

Universidad ORT Uruguay

Facultad de Ingeniería

Vehículos Eléctricos Conectados:

Diagnóstico remoto como herramienta de pronóstico
y gestión de la salud de la batería

Entregado como requisito para la obtención del título de
Ingeniero en Electrónica

Pilar Busquets - 251364

Ignacio Ortiz de Taranco - 303624

Nicolás Valdés - 229970

Tutor: Nicolás Cremona

Declaración de autoría

Nosotros, Pilar Busquets, Ignacio Ortiz de Taranco y Nicolás Valdés, declaramos que el trabajo que se presenta en esa obra es de nuestra propia mano. Podemos asegurar que:

- La obra fue producida en su totalidad mientras realizábamos el proyecto final de carrera para la obtención del título de Ingeniero en Electrónica;
- Cuando hemos consultado el trabajo publicado por otros, lo hemos atribuido con claridad;
- Cuando hemos citado obras de otros, hemos indicado las fuentes. Con excepción de estas citas, la obra es enteramente nuestra;
- En la obra, hemos acusado recibo de las ayudas recibidas;
- Cuando la obra se basa en el trabajo realizado conjuntamente con otros, hemos explicado claramente qué fue contribuido por otros, y qué fue contribuido por nosotros;
- Ninguna parte de este trabajo ha sido publicada previamente a su entrega, excepto donde se han realizado las aclaraciones correspondientes.



Pilar Busquets



Ignacio Ortiz de Taranco



Nicolás Valdés

Agradecimientos

A nuestras familias, amigos, a Nicolás Cremona, Gonzalo Díaz de Smartway, José Manuel Lestido de Julio César Lestido (Volkswagen), Germán Ferreño y Juan Manuel Carugati del equipo de Volkswagen Autos, Ricardo Cabrera de Volkswagen Camiones, Silvana Morales de Lestido Renting, Claudio Misail, UNIT (Instituto Uruguayo de Normas Técnicas), Equipo de Electromovilidad de UTE, Erika Teliz, a todos nuestros amigos.

Abstract

Este proyecto se centra en el análisis de la vida útil de vehículos eléctricos con el objetivo de examinar el estado de salud de las baterías y determinar los factores claves de su degradación. Para realizar esto, fue necesario decodificar una serie de parámetros directamente de la comunicación de vehículos y acondicionar un dispositivo que recopile estos datos en una plataforma web. Una vez obtenidos esos datos, se desarrollaron diversos modelos predictivos con herramientas de *Machine Learning* para poder determinar las variables clave en la degradación.

En este trabajo se investigó sobre las normativas que regulan a los vehículos eléctricos a nivel nacional e internacional. Se estudió a nivel físico cómo el vehículo maneja las comunicaciones internamente mediante *softwares* especializados. Se analizaron dos modelos eléctricos diferentes: un auto y un camión eléctrico, con el fin de comprender las diferencias entre ellos y la manera en la que sus baterías se degradan. Esto fue posible gracias a la colaboración de Julio César Lestido que puso a disposición del proyecto ambos ejemplares y a la empresa Smartway que brindó las herramientas y experiencia. Sin ellos no hubiera sido posible la instalación de un dispositivo IoT que reporte las tramas CAN decodificadas del vehículo para su análisis.

En este documento además se presentan todos los análisis realizados a los datos reales del vehículo obtenidos y la creación de decenas de modelos que predigan el estado de salud variando los datos de entrenamiento, los parámetros de entrada, la arquitectura del modelo y la función de costo asociada. Sumado a eso, se realiza un análisis explicativo y un análisis de correlaciones para encontrar las causas de la degradación de las baterías.

Palabras Clave

Vehículo eléctrico, Electromovilidad, Evaluación de ciclo de vida, Homologación de vehículos eléctricos, Estrategias de gestión de energía, Reducción de emisiones, Protocolo de comunicación de baterías, ISO 11898, ISO 14229, ISO 15765, SAE J1939, SAE J1979, Bus CAN (*Controller Area Network*), Trama CAN, Identificador (ID) de mensaje, Análisis de señales, Codificación y decodificación, *Sniffer CAN*, *VehicleSpy*, *ValueCAN*, ODIS, Estado de carga (SoC), Estado de salud (SoH), Sistema de gestión de batería (BMS), Análisis de degradación, *Machine Learning*, LSTM, Mejora en la precisión de predicción, Optimización de recursos, Reducción de costos, Rendimiento, Ciclo de vida de la batería, Degradación de baterías de litio, *IoT*.

Abreviaturas

- AC: *Alternating Current*. Corriente Alterna.
- BMS: *Battery Management System*. Sistema de Gestión de Batería.
- CAN: *Controller Area Network*. Red de Área de Control.
- CCS: *Combined Charging System*. Sistema de Carga Combinada.
- CHAdeMO: *Charge de Move*. Carga para el Movimiento (estándar de carga japonés).
- DC: *Direct Current*. Corriente Continua.
- DoD: *Depth of Discharge*. Profundidad de Descarga.
- ECU: *Electronic Control Unit*. Unidad de Control Electrónico.
- EIS: *Electrochemical Impedance Spectroscopy*. Espectroscopía de Impedancia Electroquímica.
- EKF: *Extended Kalman Filter*. Filtro de Kalman Extendido.
- EV: *Electric Vehicle*. Vehículo Eléctrico.
- ICE: *Internal Combustion Engine*. Motor de Combustión Interna.
- IEC: *International Electrotechnical Commission*. Comisión Electrotécnica Internacional.
- IoT: *Internet of Things*. Internet de las Cosas.
- IR: *Internal Resistance*. Resistencia Interna.

- ISO: *International Organization for Standardization*. Organización Internacional de Normalización.
- ISO-TP: *ISO Transport Protocol*. Protocolo de Transporte ISO.
- LAM: *Loss of Active Material*. Pérdida de Material Activo.
- LAMNE: *Loss of Active Material in Negative Electrode*. Pérdida de Material Activo en el Electrodo Negativo.
- LAMPE: *Loss of Active Material in Positive Electrode*. Pérdida de Material Activo en el Electrodo Positivo.
- LFP: *Lithium Iron Phosphate*. Fosfato de Hierro y Litio.
- LLI: *Loss of Lithium Inventory*. Pérdida del Inventario de Litio.
- LSTM: *Long Short-Term Memory*. Memoria a Corto y Largo Plazo.
- ML: *Machine Learning*. Aprendizaje Automático.
- NCA: *Nickel Cobalt Aluminum*. Níquel Cobalto Aluminio.
- NMC: *Nickel Manganese Cobalt*. Níquel Manganeso Cobalto.
- OBC: *On-Board Charger*. Cargador a Bordo.
- OBD: *On-Board Diagnostics*. Diagnóstico a Bordo.
- OBD-II: *On-Board Diagnostics II*. Segunda versión del sistema de diagnóstico a bordo, estandarizada en 1996.
- RUL: *Remaining Useful Life*. Vida Útil Restante.
- SAE: *Society of Automotive Engineers*. Sociedad de Ingenieros Automotrices.
- SHAP: *SHapley Additive exPlanations*. Explicaciones Aditivas de Shapley.
- SoC: *State of Charge*. Estado de Carga de la Batería.
- SoH: *State of Health*. Estado de Salud de la Batería.
- UDS: *Unified Diagnostic Services*. Servicios de Diagnóstico Unificado.
- XAI: *Explainable Artificial Intelligence*. Inteligencia Artificial Explicable.

Índice

1. Introducción	18
1.1. Auge de los vehículos eléctricos	18
1.2. Electromovilidad en Uruguay	18
1.3. Desafíos de la movilidad eléctrica	19
1.4. Motivación	20
1.5. Objetivo	21
1.6. Etapas del proyecto y Resultados clave	21
2. Investigación	23
2.1. Componentes de un vehículo eléctrico	23
2.2. ¿Que es un BMS?	25
2.3. Baterías eléctricas	27
2.3.1. Introducción a las baterías eléctricas	27
2.3.2. Evolución de baterías de litio en la industria automotriz	28
2.3.2.1. Densidad energética y su impacto en el diseño de baterías	29
2.3.2.2. Composición de las baterías de EVs	29
2.3.3. Ciclo de carga y degradación de la batería	30
2.3.4. Parámetros de la batería	32
2.3.4.1. Métodos utilizados para la estimación del SoH	34
2.3.4.2. Dificultades en el cálculo y estimación de SoH	36
2.4. Tipos de carga de vehículos eléctricos	37
2.4.1. Modos de carga y estándares internacionales	37
2.4.2. Norma IEC 62196, conectores de carga	39

2.4.2.1.	Conectores para carga en corriente alterna (AC, modos 2 y 3):	39
2.4.2.2.	Conectores de carga en corriente continua (DC, modo 4): . . .	40
2.4.3.	Módulo de comunicación entre vehículo y estación de carga	40
2.4.3.1.	Componentes involucrados	41
2.4.3.2.	Importancia en la carga segura	41
2.4.4.	Infraestructura de carga de EVs en Uruguay	42
2.4.5.	Frenado regenerativo	45
2.4.5.1.	Funcionamiento	45
2.4.5.2.	Beneficios	45
3.	Primera etapa: Obtención de datos de vehículos	46
3.1.	Evolución de los sistemas de comunicación y diagnóstico en vehículos eléctricos	46
3.2.	¿Qué es CAN?	49
3.2.1.	¿En qué consiste la capa física?	50
3.2.2.	¿Qué es una ECU?	51
3.2.3.	¿Qué es una trama?	52
3.2.4.	Versión alternativa de CAN: 29 <i>bits</i>	54
3.2.5.	Comunicación <i>multi-frame</i>	54
3.2.6.	Clasificación de CAN en función de la velocidad de transmisión . . .	56
3.3.	¿Qué es UDS?	57
3.4.	¿Qué es OBD?	58
3.5.	Decodificación de datos	59
3.5.1.	Acceso a vehículos eléctricos	59
3.5.2.	Proceso de ingeniería inversa	59
3.5.3.	<i>Hardware</i> y <i>software</i> utilizado	60

3.5.3.1.	ODIS y herramienta de diagnóstico	63
3.5.3.2.	<i>ValueCAN</i> y <i>VehicleSpy - Hardware</i> y <i>software</i> de decodificación	63
3.5.4.	Volkswagen e UP! (auto eléctrico)	65
3.5.5.	Volkswagen e Delivery (camión eléctrico)	71
3.5.5.1.	Problema 1: Mensajes <i>broadcast</i> sin decodificar	72
3.5.5.2.	Problema 2: Tercera línea CAN	73
3.5.5.3.	Conexiones para ingeniería inversa	74
3.5.6.	Comparación entre vehículos analizados	75
3.5.7.	Resumen	76
3.6.	Recopilación de datos	78
3.6.1.	Acondicionamiento del dispositivo	78
3.6.2.	Puesta a prueba del dispositivo	79
3.6.3.	Instalación del dispositivo	81
4.	Segunda etapa: Análisis de datos	83
4.1.	Metodología aplicada	83
4.2.	Exploración de los datos recopilados	84
4.2.1.	Pre-procesamiento de datos	85
4.2.2.	Depuración de los datos	86
4.2.2.1.	Ciclos de carga	86
4.2.2.2.	Ciclos de descarga	90
4.2.3.	Estimación de la vida útil de la batería	93
4.3.	Predicción de curva de degradación de SoH	98
4.3.1.	Métodos de predicción	98
4.3.2.	Conjuntos de datos utilizados	100

4.3.2.1.	Conjunto de datos de NASA	100
4.3.2.2.	Conjunto de datos de MIT	102
4.3.2.3.	Comparación de los conjuntos de datos	104
4.3.3.	Selección de arquitectura del modelo	105
4.3.4.	Introducción a LSTM	107
4.3.4.1.	Descripción de la celda LSTM	107
4.3.4.2.	Formato de entrada al modelo LSTM	108
4.3.4.3.	Descripción de la arquitectura elegida	109
4.3.5.	Pre-procesamiento de los <i>datasets</i>	110
4.3.5.1.	Definición de features	111
4.3.5.2.	Selección de <i>features</i>	113
4.3.5.3.	Normalización de features	114
4.3.5.4.	Ventana de entrada al modelo	115
4.3.5.5.	Tamaño de la ventana deslizante y horizonte temporal objetivo	116
4.3.6.	Entrenamiento del modelo	116
4.3.6.1.	Implementación	117
4.3.6.2.	CUDA	117
4.3.6.3.	cuDNN	118
4.3.6.4.	Compatibilidad y versiones del software utilizado	118
4.3.6.5.	Optimización de rendimiento mediante <i>batch size</i>	118
4.3.7.	Pruebas realizadas	119
4.3.7.1.	Modelos entrenados con el <i>dataset</i> de NASA	119
4.3.7.2.	Modelos entrenados con el <i>dataset</i> del MIT	121
4.3.8.	Evaluación del modelo sobre datos reales	123
4.3.9.	Resumen	125

4.3.10. Limitaciones del método elegido	126
4.4. Análisis de patrones de degradación	129
4.4.1. Análisis de correlación de <i>features</i>	129
4.4.2. Valores Shapley	130
4.4.2.1. Limitaciones del análisis mediante valores Shapley	132
4.4.3. Resultados de análisis de SHAP	134
4.4.3.1. <i>Global Bar Plot</i> :	134
4.4.3.2. <i>Beeswarm Plot</i> :	135
4.4.3.3. Limitaciones del análisis	139
4.4.4. Conclusiones de ambos análisis	139
5. Evaluación de la gestión de riesgos	141
5.1. Falta de estandarización	141
5.2. Acceso a vehículos eléctricos	141
5.3. Interferencias en la comunicación CAN: posible causante de fallas	142
5.4. Incertidumbre en la cantidad de datos recopilados	142
5.5. Resumen	143
6. Evolución del proyecto: Diagramas de Gantt	143
7. Análisis Final	147
7.1. Cumplimiento de objetivos	147
7.2. Conclusiones	148
7.3. Trabajo Futuro	150
7.3.1. Volkswagen e UP	150
7.3.2. Volkswagen e Delivery	151

7.3.3. Análisis de datos y predicción de vida útil de la batería	151
7.4. Potencialidades del proyecto	154
8. Referencias Bibliográficas	156
A. Anexos	164
A.1. Resultados de Volkswagen e UP (auto eléctrico)	164
A.2. Conectores necesarios para la decodificación de Volkswagen e Delivery (camión eléctrico)	169
A.3. Dispositivo SG de Smartway: TDU	170
A.4. Visualización de los datos MIT	172
A.5. Resultados modelos NASA	173
A.6. Resultados modelos MIT	174
A.6.1. MSE cómo función de pérdida	174
A.6.2. MSE ponderado cómo función de pérdida	175
A.7. Resultados modelo NASA con datos reales	176
A.8. Resultados modelo MIT con datos reales	178
A.9. Gráficos <i>beeswarm</i> del análisis de SHAP	181

Índice de cuadros

3.1. Modelo OSI aplicado a CAN y UDS	49
4.1. Tabla comparativa de modelos realizados con <i>dataset</i> de NASA.	120
7.1. Resultados de los objetivos de la etapa 1.	147
7.2. Resultados de los objetivos de la etapa 2.	148
A.1. Lista de ECUs y sus nombres en VW e UP	164
A.2. Tabla de referencia para parámetros de kilometraje y averías	164
A.3. Tabla de referencia para los estados del cargador	164
A.4. Tabla de referencia para parámetros adicionales de la batería	165
A.5. Tabla de parámetros CAN	166
A.6. Indicadores luminosos del dispositivo TDU	171

Índice de figuras

2.1. Esquema de vehículo <i>full</i> eléctrico [1]	23
2.2. Topología BMS centralizada [2]	26
2.3. Topología BMS modular [2]	26
2.4. Topología BMS Primaria / Secundaria [2]	27
2.5. Composición de batería eléctrica [3]	28
2.6. Diagrama temporal de ciclo de carga [4]	31
2.7. Diagrama de SoC, DoD y Capacidad de la batería [5]	33
2.8. Modelo eléctrico simplificado de una batería [6]	35
2.9. Modos de carga estandarizados en la norma IEC 61851-1 [7]	38
2.10. Conectores estandarizados por la IEC 62196 [8]	39
2.11. Ejemplo de SAVE instalado [9]	42
2.12. Mapa de puntos de carga en Uruguay [9]	43
2.13. Ejemplo de punto de carga	44
3.1. Conector SAE J1962, más conocido como conector OBD [10]	47
3.2. Niveles de voltaje de líneas CAN según el estado del bus [11]	50
3.3. Principales componentes de la ECU [12]	52
3.4. Estructura de las tramas de CAN estándar de 11 bits [12]	53
3.5. Comparación CAN estándar y CAN extendido [13]	54
3.6. Tramas definidas para comunicación <i>multi-frame</i> [14]	55
3.7. Estructura de los mensajes del servicio 0x22 de UDS[15]	57

3.8. Conector OBD en vehículo [16]	60
3.9. Diagrama del proceso de ingeniería inversa realizado	62
3.10. Herramienta <i>ValueCAN</i> [17]	64
3.11. Dispositivo ValueCAN y software VehicleSpy en funcionamiento	65
3.12. Mapa representativo de ECUs en Volkswagen e UP	66
3.13. Foto analizando el Volkswagen e UP	67
3.14. Diagrama de conexiones para la decodificación de parámetros del camión. . .	74
3.15. Dispositivo SG de Smartway	78
3.16. Diagrama de sistema de prueba de dispositivo SG de Smartway	80
3.17. Instalación del Dispositivo SG de Smartway en Volkswagen e UP	81
3.18. Diagrama del Dispositivo SG instalado en Volkswagen e UP	82
4.1. Porcentaje de utilización del vehículo	85
4.2. Gráficas de corriente para ciclos de carga	87
4.3. Gráficas de corriente para ciclos de carga	88
4.4. Variación de la temperatura en los ciclos de carga	89
4.5. Gráficas de corriente para ciclos de carga	90
4.6. Gráficas de tensión para ciclos de descarga	91
4.7. Gráficas de temperatura para ciclos de descarga	93
4.8. Serie temporal de capacidad en función del SOC	94
4.9. SoH vs número de ciclo	95
4.10. SoH vs número de ciclo	96
4.11. SoH vs temperatura promedio	97

4.12. B005 - Ciclo de carga	101
4.13. B005 - Ciclo de descarga	101
4.14. Evolución de SoH para cada celda	102
4.15. Gráficas de voltaje, corriente y temperatura para la Celda 14 - Ciclo 429 . . .	103
4.16. Evolución del SoH por celda - MIT	104
4.17. Cálculo del estado interno de una celda en un modelo [18]	107
4.18. Arquitectura del modelo LSTM implementado	110
4.19. Esquemático de la ventana de entrada al modelo LSTM	115
4.20. Predicciones del SoH de la batería en el vehículo intervenido utilizando el modelo NASA	124
4.21. Predicciones del SoH de la batería en el vehículo intervenido utilizando el modelo MIT	125
4.22. Gráfica de barras de <i>ensamble</i> de modelos.	134
4.23. Gráficas de <i>beeswarm</i> del análisis SHAP	137
6.1. Diagrama de Gantt Final	146
A.1. Imágenes de conectores para decodificación del camión eléctrico	169
A.2. Visualización de las señales en celdas MIT	172
A.3. Resultados de los modelos NASA	173
A.4. Resultados de los modelos MIT con función de pérdida MSE	174
A.5. Resultados de los modelos MIT (Parte 1)	175
A.5. Resultados de los modelos MIT (Parte 2)	176
A.6. Resultados de los modelos NASA aplicados a datos reales del vehículo	177
A.7. Resultados de los modelos aplicados a datos reales del vehículo (Parte 1) . . .	178

A.7. Resultados de los modelos aplicados a datos reales del vehículo (Parte 2) . . .	179
A.7. Resultados de los modelos aplicados a datos reales del vehículo (Parte 3) . . .	180
A.8. Gráficos <i>beeswarm</i> del análisis de SHAP por <i>fold</i> (Parte 1)	181
A.8. Gráficos <i>beeswarm</i> del análisis de SHAP por <i>fold</i> (Parte 2)	182

1. Introducción

1.1. Auge de los vehículos eléctricos

En las últimas décadas, la industria automotriz ha experimentado una transformación sin precedentes con la llegada y adopción masiva de los vehículos eléctricos (EV: *electric vehicle*). Esta transición ha sido impulsada por la necesidad de reducir las emisiones de gases de efecto invernadero (GEI) y disminuir la dependencia de combustibles fósiles, ya que son factores clave en la lucha contra el cambio climático y en la seguridad energética global.

Según la Agencia Internacional de Energía (IEA), el mercado de vehículos eléctricos está creciendo a un ritmo exponencial. En el primer trimestre de 2023, las ventas superaron los 2.3 millones de unidades, reflejando un incremento del 25 % en comparación con el mismo período de 2022 [19]. Este crecimiento se debe, en gran parte, a políticas de incentivo gubernamental, avances tecnológicos en baterías y una mayor concienciación sobre el impacto ambiental del transporte tradicional.

A nivel global, los actores clave en la industria incluyen fabricantes como Tesla, BYD, Volkswagen, Nissan y Rivian, así como gobiernos que han implementado subsidios, incentivos fiscales y normativas estrictas de emisiones para fomentar la adopción de EVs. En particular, la Unión Europea y China han liderado la transición hacia la electrificación, estableciendo regulaciones que restringen la venta de vehículos a combustión en las próximas décadas.

1.2. Electromovilidad en Uruguay

En Uruguay, la transición hacia la movilidad eléctrica ha sido impulsada tanto por el sector público como por el privado, con iniciativas enfocadas en facilitar la adopción de vehículos eléctricos. Algunas de las estrategias implementadas incluyen:

- **Reducción de impuestos:** Exoneración del Impuesto Específico Interno (IMESI) en la compra de vehículos eléctricos, reduciendo su costo en comparación con modelos de combustión.

- **Subsidios en la tarifa eléctrica:** La empresa estatal UTE ha desarrollado tarifas diferenciadas con bonificaciones para la carga de EVs en horarios de menor demanda energética.
- **Tarifas inteligentes:** Además UTE implementa nuevas tarifas para incentivar a los usuarios a utilizar al máximo la energía al menor costo monetario. Estas tarifas manejan precios diferenciados según la hora del día, buscando que los usuarios carguen sus vehículos en horarios de menor demanda.
- **Infraestructura de carga:** Implementación de una red de carga pública con cobertura en todo el territorio nacional, adaptada a los diferentes estándares de conectores disponibles en el mercado. Más de 300 puntos de carga han sido instalados y presentan una diversidad de conectores y potencias de carga amplia para adaptarse a las necesidades de los usuarios.
- **Iniciativas en el transporte público:** Incorporación de ómnibus eléctricos y vehículos oficiales de cero emisiones en flotas gubernamentales.

Estos esfuerzos han permitido un crecimiento sostenido en la cantidad de vehículos eléctricos en circulación, aunque persisten desafíos relacionados con la infraestructura de carga en zonas rurales y la percepción de los consumidores respecto a la autonomía y costo de estos vehículos.

1.3. Desafíos de la movilidad eléctrica

Si bien la movilidad eléctrica representa una solución clave para la reducción de emisiones, no se han considerado con la misma profundidad los desafíos asociados a la vida útil de las baterías y su impacto en el ciclo de vida del vehículo eléctrico. Los procesos más contaminantes de la electromovilidad son aquellos relacionados con la producción y el reciclaje de las baterías, que requieren una gran cantidad de energía y recursos naturales, además de perjudicar el medio ambiente relacionados a la contaminación de suelos y cuerpos de agua.

Uruguay reconoce la necesidad de actuar y ha comenzado a buscar soluciones para mitigar estos problemas. Una nota del diario “El País” entrevista a la doctora en Química Erika Teliz diciendo lo siguiente: “Cuanto más tiempo podamos usar las baterías, menos materiales necesitaremos extraer para generar nuevas” [20].

1.4. Motivación

Estas situaciones plantean preguntas cruciales sobre la sostenibilidad y la viabilidad económica de los vehículos eléctricos. Por ejemplo, ¿cuándo una batería deja de ser apta para un vehículo? ¿Cómo se define el estado de salud de una batería? ¿Cómo se conoce el estado de salud y la vida útil restante de un vehículo eléctrico usado? ¿Qué factores afectan a la degradación de estas baterías?

Pero más importante:

¿Es posible anticiparse a la degradación de la batería para ralentizarla?
--

Estas preguntas tienen implicaciones tanto ambientales como económicas. En primer lugar, el reciclaje y la reutilización de baterías aún enfrentan desafíos tecnológicos y logísticos, lo que podría generar un problema de residuos a gran escala. En segundo lugar, la degradación de la batería afecta directamente el valor de reventa de los vehículos eléctricos, influyendo en la decisión de compra de los consumidores y en la adopción masiva de esta tecnología.

A pesar de que los costos de las baterías de vehículos eléctricos disminuyen anualmente, el precio de las baterías representa hoy en día el 20 % del costo total de un vehículo eléctrico promedio, con el valor de la batería ajustado a \$139 USD/*kWh* en 2023 [21]. Esta cifra se estima considerando costos del vehículo y de la batería en su país de origen, y no incluye costos de transporte ni impuestos. En el caso de Uruguay, el costo de la batería puede ser aún mayor debido a los altos impuestos de importación y a la falta de infraestructura local para su producción.

Por lo tanto, comprender el estado de salud (SoH: *State of Health*) de la batería y desarrollar métodos eficientes para su monitoreo es fundamental para mejorar la sostenibilidad y la rentabilidad de los vehículos eléctricos en su ciclo de vida completo. Este proyecto busca abordar estas problemáticas mediante el análisis de los factores que afectan la degradación de las baterías y su impacto en la autonomía, la seguridad y el valor de reventa de los vehículos eléctricos en Uruguay.

1.5. Objetivo

Se empleó la metodología OKR (*Objectives & Key Results*) para definir los objetivos del proyecto. Este enfoque se compone de dos elementos fundamentales:

- Objetivo (O): Representa la meta a alcanzar, caracterizándose por su naturaleza cualitativa, inspiradora y ambiciosa.
- Resultados Clave (KR): Son indicadores medibles que reflejan el avance hacia el objetivo. Permiten evaluar de manera cuantitativa el progreso y determinar si la meta se ha logrado.

Este proyecto tiene como objetivo el análisis del estado de salud de baterías de vehículos eléctricos reales, examinando qué factores son los más influyentes para su evolución en el tiempo. A su vez, se busca obtener un método para poder predecir el estado de salud futuro a partir de las características de datos de conducción y carga, mediante el entrenamiento de algoritmos de *Machine Learning*.

1.6. Etapas del proyecto y Resultados clave

Para cumplir este objetivo, es posible separar el proyecto en dos grandes etapas: primero la recopilación de los datos y luego, su análisis.

1. Primera etapa: Obtención de datos de vehículos: Consistirá en comprender la comunicación CAN y el protocolo OBD-II para la decodificación de los datos del vehículo. De esta etapa, se definen como resultados clave:
 - 1.1 Investigar y comprender en profundidad el protocolo de comunicación CAN (*Controller Area Network*), el sistema OBD-II (*On-Board Diagnostics*), las ECU (*Electronic Control Unit*), el sistema BMS (*Battery Management System*) y todo material adicional que resulte relevante para el proyecto.
 - 1.2 Decodificar parámetros de dos vehículos eléctricos diferentes. Esto implica localizar los parámetros de interés para el análisis dentro de las unidades de control de los vehículos e interpretarlos. Una vez conocida la estructura para pedir estos datos, interpretar la respuesta y registrarla en una base de datos.

- 1.3 Posterior a la generación de la base de datos, es vital reportar los mismos a una base de datos externa para comenzar con la segunda etapa del proyecto.
2. Segunda etapa: Análisis de datos: Con los datos obtenidos, se planea entrenar modelos para predecir el estado de salud. Los resultados clave en esta etapa son los siguientes:
 - 2.1 Analizar los datos recopilados, conocer sus distribuciones y realizar un primer cálculo o estimación del estado de salud actual.
 - 2.2 Investigación del estado del arte de las arquitecturas de Machine Learning en cuanto a la predicción de series temporales. Una vez realizado esto, entrenar y evaluar los modelos predictivos obtenidos, comparándolos con métricas relevantes.
 - 2.3 Obtener un análisis en profundidad sobre qué aspectos afectan mayoritariamente a la degradación de la batería en el EV. El enfoque de el mismo será desarrollar estrategias para mitigar o ralentizar la degradación.

En resumen, en la primera etapa se busca comprender cómo los vehículos eléctricos se comunican internamente, para decodificar los parámetros a analizar en la segunda etapa. Por otra parte, la segunda etapa se centra en el estudio de los datos recopilados, conocer sus distribuciones y generar modelos predictivos que permitan estimar el estado de salud de la batería al mismo tiempo que se analizan los factores que afectan su degradación.

En este documento se verán las adversidades encontradas en el transcurso del proyecto y cómo estas fueron mitigadas y superadas. Se presentarán los resultados obtenidos de los análisis realizados, identificando qué modelos (o qué conjunto de hiperparámetros) obtuvieron menores índices de error en sus predicciones. Por último, se evaluarán los efectos de los parámetros estudiados sobre el estado de salud de la batería, en búsqueda de estrategias de ralentización de la degradación de la batería.

2. Investigación

2.1. Componentes de un vehículo eléctrico

Antes de comenzar a estudiar sobre el estado de salud, fue necesaria una investigación sobre la composición de un vehículo eléctrico para entender los sistemas internos y, posteriormente, analizarlos y estudiarlos en mayor profundidad.

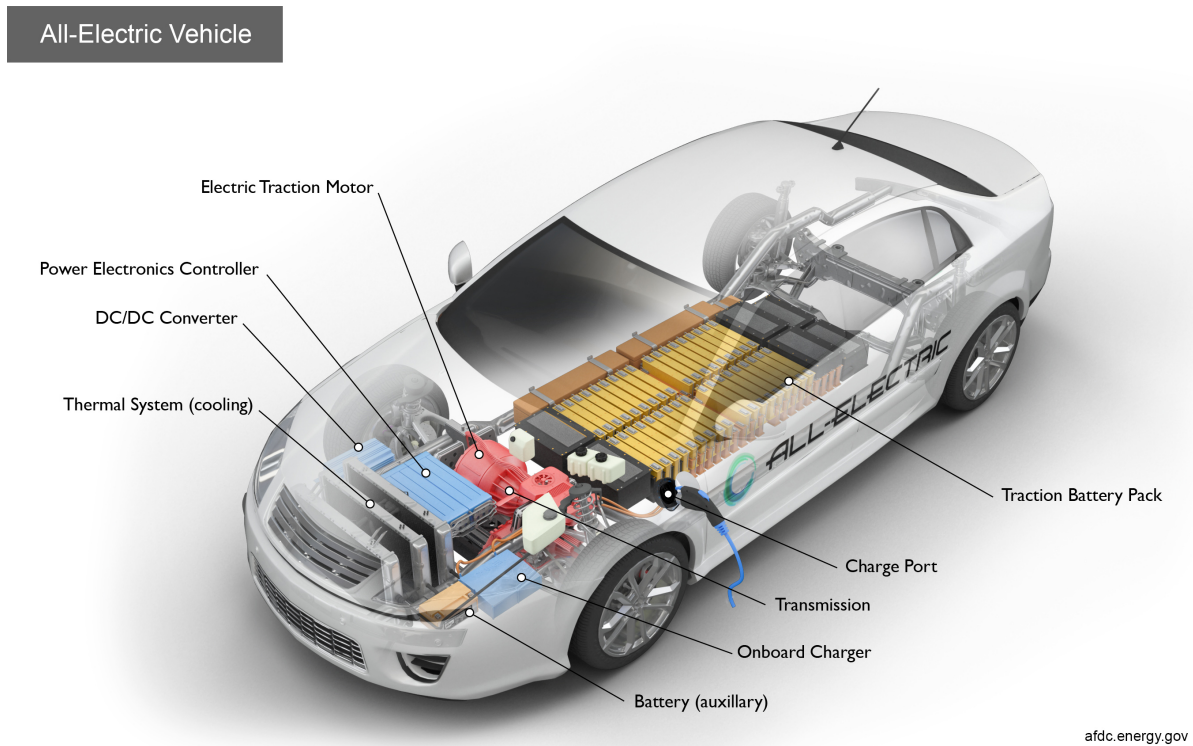


Figura 2.1: Esquema de vehículo *full* eléctrico [1]

En la figura 2.1 se observan los componentes principales de este tipo de vehículos: [1]

- **Motor de tracción eléctrica (motor eléctrico):** Encargado de convertir la energía eléctrica almacenada en la batería en movimiento mecánico para la propulsión del vehículo. Dependiendo del vehículo, existen motores de corriente continua y de corriente alterna.
- **Puerto de carga:** Proporciona el conector para la entrada de corriente continua o alterna para la carga de la batería.

- **Paquete de baterías de tracción (batería):** Compuesto de celdas de iones de litio (a ampliar en la sección posterior) y utilizado para el almacenamiento de energía eléctrica y la alimentación del motor eléctrico.
- **Cargador a bordo (OBC: *On-Board Charger*):** Encargado de convertir la corriente alterna (CA) de la red eléctrica en corriente continua (CC) para cargar la batería. El OBC no actúa cuando la energía entrante es CC, como en el caso de los cargadores rápidos. También puede convertir corriente continua en alterna para devolver energía a la red y otros dispositivos eléctricos, como a hogares y a otros vehículos eléctricos [22] [23] [24]. Dependiendo del vehículo, la potencia máxima que el OBC puede entregar en forma de corriente continua a la batería para realizar la carga varía entre 3,6 kW y 22 kW.
- **Controlador de electrónica de potencia:** Regula la energía entregada al motor y gestiona la aceleración y rendimiento del vehículo.
- **Convertidor DC/DC:** Reduce el voltaje para alimentar los sistemas auxiliares: pantallas, iluminación, sistemas de confort, entre otros.
- **Batería auxiliar:** Suministra energía a los sistemas electrónicos del vehículo: radio, computadora a bordo, etc.
- **Transmisión:** En el caso de un motor de combustión interna, este genera potencia en un rango limitado de revoluciones por minuto y con una caja de cambios es posible ajustar la relación entre el motor y las ruedas para aumentar o disminuir el torque. En EVs, los motores eléctricos entregan torque instantáneo, lo que elimina la necesidad de la caja de cambios. En lugar de cambiar marchas, una relación de reducción fija disminuye la velocidad y aumenta el torque para un movimiento más eficiente.
- **Sistema de gestión térmica o Sistema de gestión de batería (BMS: *Battery Management System*):** Sistema electrónico que monitorea y gestiona la batería para garantizar seguridad, rendimiento y longevidad. Es el principal encargado de regular la potencia en el caso de carga con corriente continua (en el caso de corriente alterna, la potencia queda limitada por el OBC). Dentro de sus funciones se encuentra: [25]
 - **Monitoreo de celdas de litio:** Supervisión de tensión, corriente y temperatura de cada celda de la batería para evitar sobrecargas o descargas profundas.
 - **Balanceo de celdas:** Asegura que todas las celdas mantengan un estado de carga uniforme, prolongando la vida útil del mismo.

- **Protección térmica:** Coordina los sistemas de refrigeración y calefacción.
- **Estimación del estado de salud:** Si bien lo calcula, este dato no fue accesible mediante la herramienta de diagnóstico utilizada (más adelante se detalla cómo se obtuvieron los datos).
- **Gestión de energía:** Controla la energía proporcionada al motor desde la batería y la recarga durante la regeneración del frenado.

2.2. ¿Que es un BMS?

Los vehículos modernos, especialmente los eléctricos, dependen de múltiples *Electronic Control Units* (ECUs) para gestionar distintos sistemas, desde el tren motriz hasta la seguridad y el confort. Entre estas unidades, una de las más críticas es el *Battery Management System* (BMS), un sistema electrónico diseñado para supervisar, controlar y optimizar el rendimiento de un paquete de baterías recargables, garantizando su operación segura y eficiente.

El BMS monitorea constantemente variables clave como el voltaje, la corriente y la temperatura de cada celda de la batería. No solo recopila estos datos, sino que también toma decisiones en tiempo real para evitar condiciones adversas como sobretensiones, sobrecorrientes o sobrecalentamiento de las celdas. Además, equilibra la carga entre celdas para prevenir desgastes desiguales, maximizando la vida útil y la autonomía del vehículo.

Este sistema no opera de manera aislada, sino que se comunica con otras ECUs del vehículo, como la del tren motriz y la de gestión térmica, para coordinar la entrega de energía de manera eficiente. Gracias a esta integración, el BMS permite funciones avanzadas como la regeneración de energía en frenado, la optimización del consumo según las condiciones de manejo y la detección temprana de fallos eléctricos. En conjunto, el BMS y las ECUs forman un ecosistema inteligente que garantiza la seguridad, eficiencia y longevidad de los vehículos eléctricos.

El BMS puede estar compuesto por una o más ECUs, dependiendo de la complejidad del vehículo. En algunos casos, se divide en unidades especializadas, como una ECU dedicada al monitoreo y protección de la batería y otra encargada de la gestión energética y la comunicación con el resto del vehículo. Gracias a esta otra integración, el BMS permite funciones avanzadas como la regeneración de energía en frenado, la optimización del consumo según las condiciones de manejo y la detección temprana de fallos eléctricos.

Existen distintas topologías de BMS:

- **BMS centralizada:** un único módulo se encarga de monitorear todas las celdas de la batería. Tiende a ser más compacto y mas económico, pero requiere de más puertos y resulta en cableados más complejos.

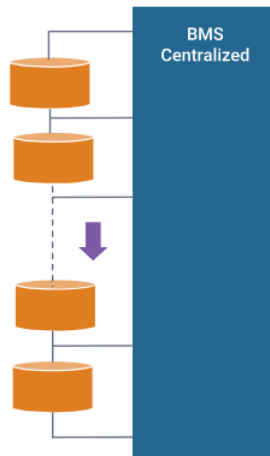


Figura 2.2: Topología BMS centralizada [2]

- **BMS modular:** Dos o más módulos, cada uno dedicado a un grupo de conexiones y celdas. Es más escalable y mantenible, pero suele ser mas caro y tener funcionalidades redundantes.

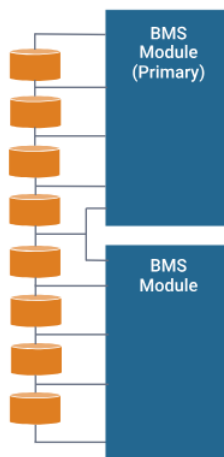


Figura 2.3: Topología BMS modular [2]

- **BMS Primario / Subordinado:** Similar a la modular, pero con un módulo primario que se encarga del control y de la comunicación externa, mientras que los secundarios son más simples y se encargan principalmente de la medición de valores.

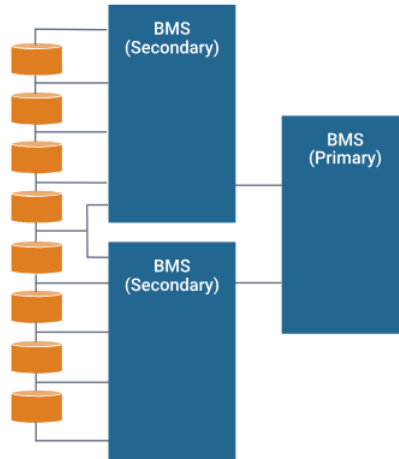


Figura 2.4: Topología BMS Primaria / Secundaria [2]

2.3. Baterías eléctricas

Fue necesario realizar un estudio más profundo para comprender el concepto del estado de salud de la batería. En esta sección se responderán preguntas como: ¿Cómo está compuesta la batería de un vehículo eléctrico? ¿Existen varios tipos de baterías en EVs y qué diferencias presentan? ¿Cómo se produce el efecto de degradación de una batería? ¿Qué efectos produce la degradación? ¿Afecta solamente a la cantidad de energía capaz de almacenar?

2.3.1. Introducción a las baterías eléctricas

El funcionamiento de una batería de un vehículo eléctrico es muy similar a la batería de un teléfono celular. Están compuestas de celdas como la que se muestra en la Figura 2.5 y dentro de cada celda se encuentran los elementos indicados en la figura. Cuando a estas celdas se les aplica una diferencia de potencial, los electrones comienzan a circular desde el ánodo hacia el cátodo a través del electrolito, mientras que los iones lo hacen en sentido contrario, creando así corriente eléctrica en la batería [3] [26]. Esto aplica tanto para cuando la batería está siendo cargada como para cuando está siendo utilizada.

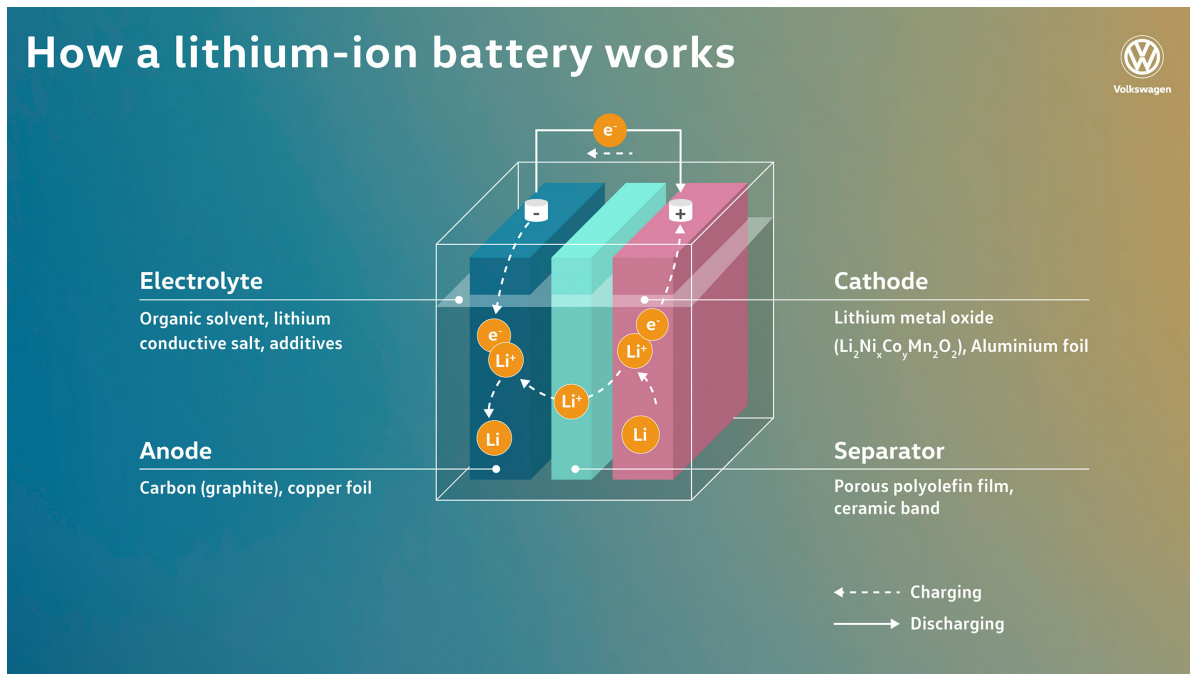


Figura 2.5: Composición de batería eléctrica [3]

2.3.2. Evolución de baterías de litio en la industria automotriz

Las baterías de iones de litio han revolucionado el almacenamiento de energía en múltiples industrias, desde la electrónica de consumo hasta la movilidad eléctrica. Su desarrollo ha estado marcado por avances en capacidad de almacenamiento, seguridad, eficiencia y materiales.

Los primeros intentos de utilizar litio en baterías se remontan a la década de 1970, cuando *Stanley Whittingham* propuso el uso de este metal en baterías recargables. Posteriormente, *John Goodenough* descubrió que el óxido de cobalto podía mejorar significativamente la capacidad y estabilidad de la celda. Sin embargo, fue *Akira Yoshino*, en 1985, quien logró desarrollar la primera batería de iones de litio segura y comercialmente viable, lo que llevó a *Sony* a introducirlas en el mercado en 1991 [27].

Inicialmente, estas baterías se utilizaron en dispositivos electrónicos portátiles, como teléfonos celulares y computadoras portátiles, ya que ofrecían una mayor autonomía en comparación con tecnologías previas como las baterías de níquel-cadmio (*Ni – Cd*) o níquel-hidruro metálico (*NiMH*). Sin embargo, con el auge de la movilidad eléctrica, la industria automotriz comenzó a adaptar y mejorar esta tecnología para vehículos eléctricos (EVs).

2.3.2.1 Densidad energética y su impacto en el diseño de baterías

Uno de los principales retos en la evolución de las baterías ha sido aumentar su densidad energética, es decir, la cantidad de energía almacenada por unidad de masa (Wh/kg) o volumen (Wh/L). Este parámetro es crucial tanto en la electrónica de consumo como en los vehículos eléctricos. Se presenta una lista a modo de comparación de densidades energéticas a lo largo del tiempo:

- Baterías de *NiMH* utilizadas en primeros vehículos híbridos: Densidad entre 60 Wh/kg y 120 Wh/kg.
- Baterías de iones de litio de celulares y EVs modernos: Densidad entre 100 Wh/kg y 265 Wh/kg.
- Baterías de fosfato de hierro y litio (*LiFePO₄*): Es un subtipo de baterías de litio y son reconocidos por su seguridad y su larga vida útil. Se utilizan ampliamente en vehículos eléctricos y en el almacenamiento de energía solar. Manejan densidades entre 90 Wh/kg y 130Wh/kg [28].
- Baterías de estado sólido (tecnología emergente): Son baterías en las cuales su electrolito es sólido en vez de líquido como en los ejemplos anteriores. No sólo reduce los riesgos de seguridad, sino que teóricamente permite hasta 2,5 veces más densidad energética que las baterías de iones de litio [28].

Baterías con mayor densidad energética permiten el desarrollo de baterías más pequeñas manteniendo el mismo rendimiento y eficiencia. Esto fue un beneficio enorme no sólo para la industria telefónica, sino que también para la industria automotriz, permitiendo el desarrollo de vehículos más livianos y con niveles de autonomía nunca antes vistos en este tipo de vehículos.

2.3.2.2 Composición de las baterías de EVs

El avance en la densidad energética y seguridad de las baterías de iones de litio ha sido posible gracias a la evolución de los materiales activos utilizados en el ánodo y el cátodo. Se presentan aquí las composiciones más utilizadas en la industria automotriz: [29] [30]

1. **Litio-Ferrofosfato (*LFP* o $LiFePO_4$):** Este tipo de batería se caracteriza por su bajo costo (80 - 100 USD/kWh), gran estabilidad térmica y alta vida útil, con una estimación de entre 2500 y 5000 ciclos de carga y descarga. Al no incluir cobalto ni níquel, su impacto ambiental y económico es menor. A pesar de su menor densidad energética (entre 90 y 160 Wh/kg), su seguridad y durabilidad la convierten en una opción preferida para vehículos de uso urbano, transporte público y flotas de reparto. Tesla ha incorporado esta tecnología en los modelos 3 e Y, así como BYD en sus modelos Han y Tang [31].
2. **Níquel-Manganeso-Cobalto (*NMC*):** Representa una solución intermedia en términos de costo (100 - 120 USD/kWh), autonomía y seguridad. Su densidad energética se encuentra entre 150 y 220 Wh/kg, y su vida útil estimada varía entre 1500 y 2000 ciclos. Es una de las químicas más utilizadas a nivel mundial debido a su buen balance entre rendimiento y estabilidad. La proporción entre los tres elementos puede ajustarse para mejorar ciertas propiedades específicas del sistema. El modelo Volkswagen ID.4 incorpora esta tecnología [32].
3. **Níquel-Cobalto-Aluminio (*NCA*):** Se destaca por ofrecer la mayor densidad energética (entre 200 y 260 Wh/kg), lo que permite una mayor autonomía por unidad de peso. Sin embargo, presenta una menor estabilidad térmica y un mayor costo de producción, estimado entre 110 y 130 USD/kWh. Su vida útil ronda los 1000 a 2000 ciclos, lo que la posiciona por debajo de alternativas como *LFP* en términos de durabilidad. Es común en vehículos de alta gama o de altas prestaciones; modelos como el Tesla S y Tesla X incorporan esta tecnología en sus baterías [33].

2.3.3. Ciclo de carga y degradación de la batería

Al igual que en los EVs, los celulares han pasado de utilizar baterías con alto contenido de cobalto (NMC/NCA) a explorar alternativas más seguras y económicas como LFP, especialmente en modelos de gama baja o dispositivos con necesidades de carga más sostenibles. Sin embargo, independientemente de la composición química, la vida útil de una batería está determinada no solo por su química, sino también por la cantidad de ciclos de carga y descarga a los que es sometida a lo largo del tiempo. Cada carga parcial y cada descarga afectan progresivamente su capacidad, dando lugar a un fenómeno conocido como degradación de la batería. Para comprender este proceso, es necesario definir primero el concepto de ciclo de carga y analizar cómo influye en el rendimiento y longevidad de la batería.

Apple define que un ciclo de carga se completa cuando se consume un 100 % de la capacidad de la batería [4], aunque este consumo no tiene por qué ser luego de una única carga. Observando el diagrama de la figura 2.6 se observa que el ciclo comienza, luego la batería se descarga un 75 %, se carga un poco y luego se consume un 25 % más para completarse. No es necesario cargar la batería a 100 % en todo momento, esto es sólo parte del ejemplo.

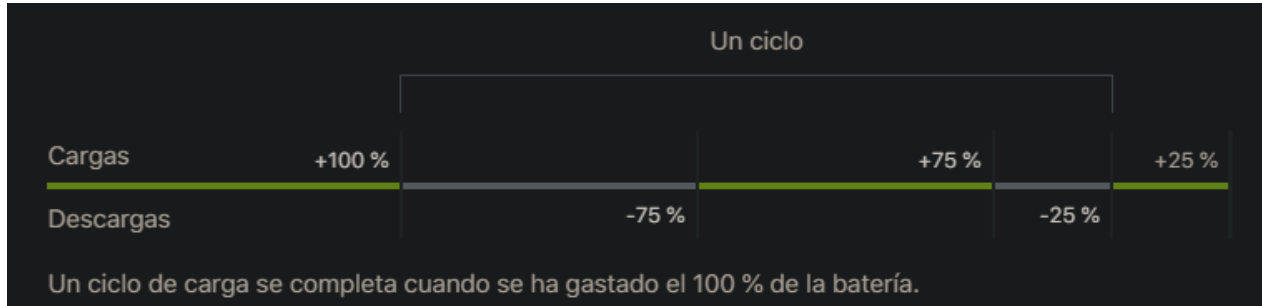


Figura 2.6: Diagrama temporal de ciclo de carga [4]

Durante el ciclo de carga circulan corrientes eléctricas que producen pequeñas degradaciones en los electrodos (ánodo y cátodo) y en el electrolito. Un estudio de *"The Ohio State University"* en *Columbus, Ohio* junto a la empresa *Hyundai Motor Company* [34] explica los mecanismos de degradación de una batería en tres modos: *"loss of lithium inventory (LLI), loss of active material (LAM), and impedance increase. LAM is further categorized into loss of active material in the anode (LAM_{NE}), and loss of active material in the cathode (LAM_{PE})"*.

La pérdida del inventario de litio (*LLI*) abarca la descomposición del litio debido a reacciones químicas internas y condiciones operativas. Esto puede causar una mayor resistencia interna y generación de calor, lo que impacta directamente la eficiencia y, sobre todo, la seguridad de la batería. La pérdida de material activo (*LAM*), por otro lado, implica que los iones de litio quedan atrapados en subproductos químicos irreversibles. Esto causa que haya menor cantidad de litio disponible para actuar y reaccionar, implicando así la disminución en la capacidad de la batería. Otros efectos adversos causados por la degradación de la batería (sin estar atado a un modo de degradación particular) son, por ejemplo: la reducción en la entrega de potencia, la ralentización en la carga rápida y mayor riesgo de fallo térmico [34]. Dado que la degradación en condiciones reales de uso está influenciada por una combinación de factores difíciles de replicar en un entorno controlado, la predicción del estado de salud de la batería no solo es un reto técnico, sino también un problema abierto en la industria de la electromovilidad.

2.3.4. Parámetros de la batería

Para evaluar el estado y rendimiento de una batería en un EV, se utilizan diversos parámetros que permiten cuantificar su capacidad y eficiencia. Un estudio conformado por varios institutos franceses [35], en conjunto con unidades didácticas de un curso de la Facultad de Ingeniería de la Universidad de la República (UdelaR) [5], explica que existen dos tipos de datos que representan la evolución de la batería: "*historical features, HF*" y las "*time series, TS*". La diferencia en estos es que los datos históricos son constantes durante todo un ciclo, mientras que la serie de datos se puede representar como una función variable en el tiempo del ciclo de carga y descarga. Dentro de los HF se encuentran, por ejemplo, el estado de salud (SoH), la vida útil restante (RUL) y la resistencia interna (IR) y en los TS encontramos datos de voltaje, corriente, temperatura, estado de carga (*SoC, state of charge*), profundidad de descarga (*DoD, depth of discharge*), entre otros.

Dentro de los TS, se destacan SoC y DoD como indicadores del estado instantáneo de la batería, que permiten conocer cuánta energía tiene almacenada la batería en este momento o cuánta energía se ha descargado. Estos indican el tiempo restante de uso hasta la próxima descarga completa y la suma de estas funciones resulta en el valor de la capacidad actual de la batería:

$$SoC(t) = \frac{C_o - \int_0^t i(t) dt}{C_o} \times 100 \%$$

$$Capacity = SoC(t) + DoD(t)$$

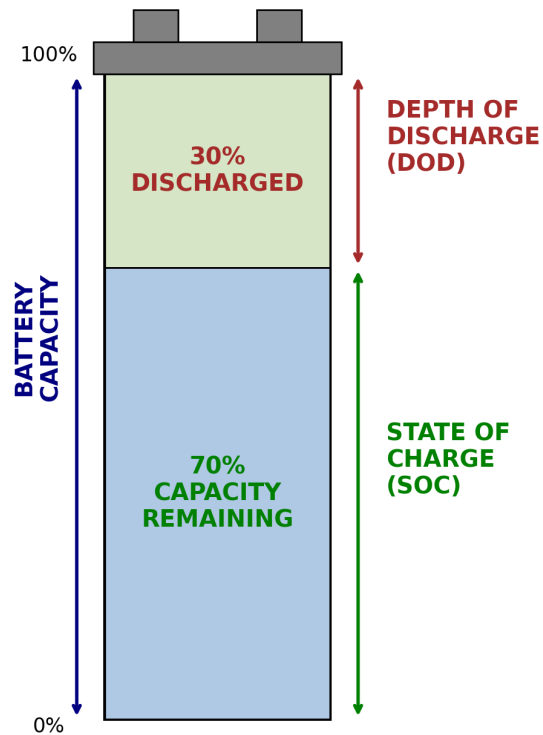


Figura 2.7: Diagrama de SoC, DoD y Capacidad de la batería [5]

1. SoC(t) se obtiene contabilizando la cantidad de carga extraída durante el período de uso (ciclo completo idealmente), integrando la corriente en el tiempo.
2. DoD(t) es la cantidad de carga eléctrica extraída durante la descarga, a partir del estado inicial (100 % de carga).

Por otra parte, dentro de los HF se tienen el SoH y RUL (*Remaining Useful Time*) y ambos parámetros estiman la vida útil restante de una batería.

1. El estado de salud se define como la cantidad máxima de carga eléctrica que puede ser extraída en el ciclo j de carga, en comparación a la cantidad de carga máxima original. Este parámetro suele expresarse en porcentaje de la capacidad inicial de la batería (C_0).

$$SoH(t) = \frac{C_{max}^j}{C_{max}^0} \times 100 \%$$

2. La vida útil restante se define como el tiempo de operación (en ciclos de carga) durante el cuál, el SoH se mantiene por encima del umbral permitido. No hay un umbral exacto definido pero la mayoría indican 70 % de la capacidad original.

2.3.4.1 Métodos utilizados para la estimación del SoH

Para la estimación del *State of Health* (SoH) de una batería, se identificaron diversos métodos basados en el análisis de parámetros observables, como la tensión, la corriente y la temperatura. Estos métodos permiten transformar estos parámetros en indicadores del estado de la batería mediante distintas estrategias de procesamiento y modelado. A continuación, se presentan los enfoques más utilizados, obtenidos a partir de la literatura consultada [6]:

1. Conteo de Coulomb

Se basa en la medición de la carga acumulada durante los ciclos de carga y descarga para estimar la capacidad disponible de la batería.

$$\left\{ \begin{array}{l} Q_{\text{discharge}} = \int_0^T I(t) dt \\ SoH [\%] = \frac{Q_{\text{discharge}}}{Q_{\text{rated}}} \times 100 \% \end{array} \right.$$

Este método es ampliamente utilizado debido a su simplicidad y bajo costo de implementación. Sin embargo, su precisión depende de la exactitud de las mediciones de corriente y del SoC inicial.

2. Método de voltaje en circuito abierto (OCV)

Este enfoque establece una relación entre el voltaje en circuito abierto y el estado de salud de la batería. Hace uso del modelo simplificado de la batería, expresando el SoH como función del voltaje en circuito abierto.

$$U_{ocv} = U + IR$$

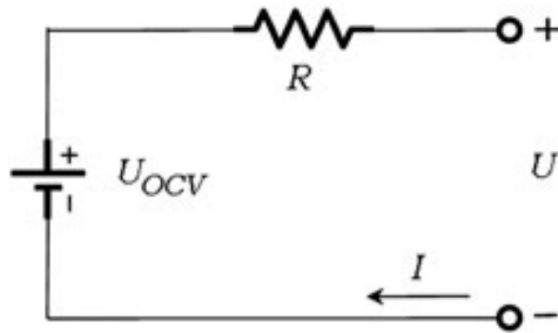


Figura 2.8: Modelo eléctrico simplificado de una batería [6]

Puede aplicarse de manera offline, mediante pruebas de laboratorio para caracterizar la curva de OCV, o de manera online, estimando el OCV a partir de mediciones en condiciones de operación. Si bien ofrece una estimación confiable, su aplicación en sistemas en tiempo real puede verse limitada debido a la dificultad de determinar el parámetro de OCV a efectos prácticos.

3. Espectroscopía de impedancia electroquímica (EIS)

Permite caracterizar la impedancia interna de la batería en función de la frecuencia, lo que posibilita la detección de variaciones asociadas al envejecimiento. Este método es altamente preciso y proporciona información detallada sobre el estado de la batería, pero requiere equipos de medición especializados y suele aplicarse en entornos de laboratorio.

4. Filtros de Kalman

Se emplean para estimar el SoH a partir de modelos dinámicos de la batería, combinando predicciones teóricas con mediciones reales. Existen variantes avanzadas, como el *Extended Kalman Filter* (EKF) y el *Unscented Kalman Filter* (UKF), que permiten mejorar la precisión en sistemas no lineales. Estos métodos presentan ventajas en términos de adaptabilidad y robustez, aunque requieren un modelado preciso y una alta capacidad de cómputo.

5. Métodos basados en aprendizaje automático

Utilizan grandes volúmenes de datos históricos para generar modelos predictivos del estado de la batería. Estos enfoques han demostrado un alto grado de precisión y capacidad de generalización, especialmente cuando se combinan con técnicas como *Relevance Vector Machines* (RVM) y redes neuronales. Sin embargo, requieren bases de datos extensas para su entrenamiento y pueden implicar un alto costo computacional.

2.3.4.2 Dificultades en el cálculo y estimación de SoH

Los métodos de estimación de SoH presentados varían de modelos más simples que requieren de bajo poder computacional hasta los más complejos que pueden incurrir en grandes costos de cómputo. Sin embargo, gran parte del error en las estimaciones que aplica para todos los métodos descritos se encuentra en la obtención de los datos. El estado de salud no es un dato que pueda medirse directamente con un sensor; implica el control de diversas variables, como el voltaje, la corriente y la temperatura. Estas variables en un entorno de laboratorio suelen ser controladas; sin embargo, suele suceder que estos datos no replican con exactitud el uso que puede tener una batería eléctrica instalada en un vehículo. El uso y la degradación de la batería de un EV se ven fuertemente influenciados por: el perfil de conducción de los choferes, el uso del frenado regenerativo, las condiciones climáticas cambiantes, el uso de "la carga rápida", el tiempo de inactividad que pueda tener.

Los métodos más simples, aquellos que emplean circuitos equivalentes y ecuaciones matemáticas, de base suelen tener mayor error en la estimación del estado de salud. Agregar "irregularidades" o factores externos como los mencionados suelen amplificar el error en la estimación. El método de la espectroscopía, como se menciona, requiere de equipos de medición especializados, lo que implicaría mayores costos de producción y mantenimiento en caso de fallas o averías.

Por otro lado se encuentran los modelos más complejos (Kalman y modelos AI) que presentan excelentes métricas de error y demuestran grandes capacidades para capturar relaciones no lineales con exactitud. Sin embargo, ambos métodos precisan de grandes volúmenes de información y datos para analizar y buscar patrones no lineales.

En conclusión, se presentan dos grandes limitaciones en la estimación del estado de salud de una batería eléctrica instalada en un EV:

1. La falta de estándares universales tanto en la obtención de los datos como en las técnicas de estimación.
2. La falta de registros de voltaje, corriente y temperatura. Si todos los fabricantes comenzaran a llevar registro de estas variables podrían generar modelos de degradación según cada modelo de vehículo. Sería hasta beneficiario para ellos mismos, permitiría tener mucha más información para un diagnóstico general y remoto.

De acuerdo con los concesionarios consultados, no se han encontrado ejemplos de fabricantes almacenando estos registros por auto. Algunos fabricantes, como BYD, están trabajando en su propia plataforma de visualización de datos, sin embargo no se prevee que esta sea de acceso público. Por lo tanto, uno de los principales desafíos de este proyecto es identificar, acceder y recopilar los datos pertinentes de los vehículos para su posterior análisis.

2.4. Tipos de carga de vehículos eléctricos

El proceso de carga en un vehículo eléctrico es un aspecto fundamental que afecta directamente su autonomía, eficiencia y durabilidad de la batería. A diferencia de los vehículos con motores de combustión interna, donde la carga de combustible es un proceso estandarizado en cuanto a tiempo y método, en los vehículos eléctricos existen diversas formas de carga que varían en velocidad, infraestructura y compatibilidad.

Con el crecimiento de la electromovilidad, se han desarrollado estándares internacionales que regulan los modos de carga, los tipos de conectores, las potencias de carga disponibles y las tecnologías de optimización energética. Estas regulaciones buscan garantizar la seguridad del usuario, la interoperabilidad entre diferentes marcas de vehículos e infraestructuras de carga y, en última instancia, la sostenibilidad del sistema eléctrico.

2.4.1. Modos de carga y estándares internacionales

Para asegurar una carga segura y eficiente, la norma IEC 61851-1 clasifica los métodos de carga de los vehículos eléctricos en cuatro modos principales, diferenciados por la potencia suministrada, la velocidad de carga y los requisitos de infraestructura. Estos modos van desde la carga lenta en tomas domésticas hasta la carga ultrarrápida en estaciones especializadas [36]:

1. **Modo 1:** Carga lenta en corriente alterna (AC) utilizando corriente doméstica estándar, sin comunicación específica entre cargador y vehículo. Debido a las limitaciones en tomas de corriente estándares, este tipo de carga suele manejar potencias hasta $3,7kW$, conocido como "carga lenta".
2. **Modo 2:** Similar al Modo 1 pero con un dispositivo adicional de control para mayor seguridad durante la carga. Maneja el mismo valor máximo de potencia entregada que el Modo 1.
3. **Modo 3:** Carga en corriente alterna utilizando una estación de carga específica capaz de comunicarse con el vehículo a cargar. Ofrece mayor seguridad y mayor control sobre la carga. Según la infraestructura instalada y la capacidad del vehículo, este modo de carga puede entregar potencias de "carga semirrápida" (entre $7kW$ y $22kW$).
4. **Modo 4:** Carga rápida utilizando corriente continua (DC) mediante una estación de carga externa encargada de rectificar la corriente alterna en corriente continua. Este modo de carga fue diseñado para "cargas rápidas" (superando los $22kW$, permitiendo cargar el 80% de la batería en 30 minutos) y para "cargas ultrarrápidas", permitiendo alcanzar los $350kW$ si el vehículo lo permite.

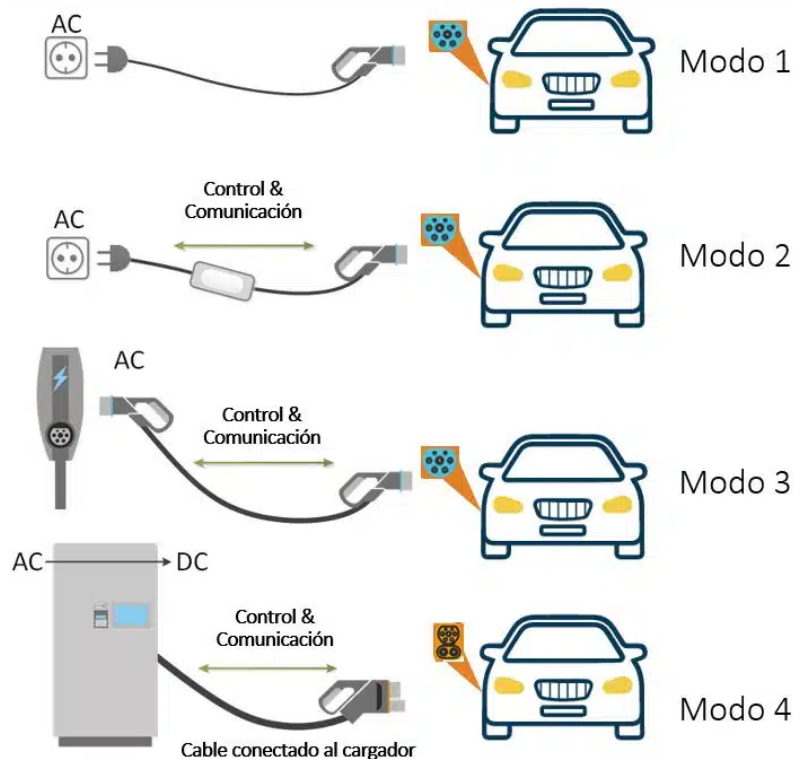


Figura 2.9: Modos de carga estandarizados en la norma IEC 61851-1 [7]

2.4.2. Norma IEC 62196, conectores de carga

Si bien la norma IEC 61851-1 establece los modos de carga en función de la potencia y el tipo de corriente utilizada, la compatibilidad y eficiencia del proceso de carga dependen en gran medida de los conectores empleados. En este sentido, la norma IEC 62196 define los tipos de conectores y enchufes utilizados en los vehículos eléctricos y estaciones de carga, asegurando la interoperabilidad entre diferentes fabricantes y mercados. A continuación, se presentan los principales conectores normalizados y su relación con los modos de carga previamente descritos.

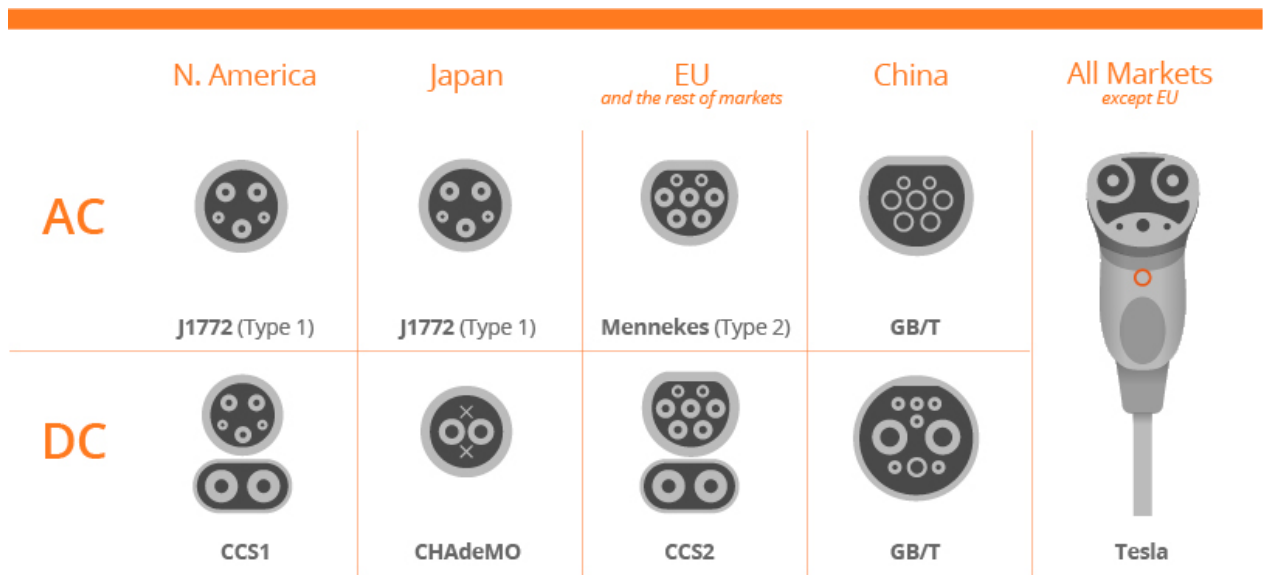


Figura 2.10: Conectores estandarizados por la IEC 62196 [8]

2.4.2.1 Conectores para carga en corriente alterna (AC, modos 2 y 3):

1. **Tipo 1 (SAE J1772):** Utilizado principalmente en América del Norte y Japón y soporta hasta $7,4kW$ **monofásico**, $32A$ a $230V$.
2. **Tipo 2 (IEC Mennekes):** Estándar europeo y sudamericano y permite soportar hasta $43kW$ **trifásico**, $63A$ a $400V$.
3. **GB/T AC:** Estándar chino que presenta un formato similar al Mennekes. Permite soportar hasta $250V$ monofásico o $440V$ trifásico, con una potencia de carga de hasta $43kW$.

2.4.2.2 Conectores de carga en corriente continua (DC, modo 4):

1. **CHAdeMO:** Estándar japonés desarrollado por Nissan y Mitsubishi, capaz de soportar hasta $400kW$ ($600A$ a $1000V$ DC).
2. **CCS:** Soporta hasta $350kW$ para cargas ultrarrápidas y es el estándar adoptado por la mayoría de fabricantes en Europa y América. Tiene dos variantes
 - 2.1 CCS1: Basado en conector Tipo 1, para Estados Unidos y Canadá.
 - 2.2 CCS2: Basado en conector Mennekes (Tipo 2), para Europa y Sudamérica.
3. **GB/T DC:** Puede alcanzar potencias de hasta $900V$ y $600A$, permitiendo cargas ultrarrápidas de $350kW$.

2.4.3. Módulo de comunicación entre vehículo y estación de carga

A partir del Modo 2, la norma IEC 61851-1 incorpora un sistema de señalización entre el vehículo y el sistema de carga, cuya función es garantizar una operación segura y adaptativa. En este modo, la señal de control proviene de una caja de control integrada en el cable (IC-CPD), que comunica al vehículo si puede iniciar la carga y bajo qué condiciones (por ejemplo, límite de corriente, fallas térmicas o desconexiones).

A partir del Modo 3 de carga definido por la norma IEC 61851-1, la interacción entre la estación de carga (EVSE) y el vehículo eléctrico (EV) requiere de un canal de comunicación activo. Esta comunicación es imprescindible para garantizar una carga segura, eficiente y conforme a las capacidades técnicas del vehículo y del sistema eléctrico. El componente responsable de esta tarea es el llamado Módulo de Control de Carga (Control Pilot), que gestiona la señalización entre ambos extremos.

Este sistema se basa principalmente en una línea de comunicación analógica de tipo PWM (modulación por ancho de pulso), transmitida a través del pin “Control Pilot” (CP) del conector, definido en la norma IEC 62196. Su propósito es intercambiar información básica como:

- Estado de conexión física y eléctrica.
- Capacidad máxima de corriente que puede entregar la estación.

- Disponibilidad para iniciar o detener la carga.
- Solicitudes del vehículo como activar la ventilación externa.

El protocolo también contempla un segundo canal opcional llamado “Proximity Pilot” (PP), que permite detectar si el cable está conectado y limitar la corriente según el tipo de cable (por ejemplo, si es monofásico o trifásico, y su sección). Este segundo canal funciona por resistencias predefinidas entre pines, sin necesidad de señal PWM.

2.4.3.1 Componentes involucrados

Los módulos de comunicación están integrados tanto en el vehículo como en la estación, y normalmente comprenden:

- Un microcontrolador encargado de interpretar y generar las señales PWM.
- Circuitos de acoplamiento y protección (aislamiento galvánico, optoacopladores, filtros).
- Drivers de señalización, incluyendo circuitos pull-up y pull-down para detectar los niveles lógicos de las señales CP y PP.
- En el caso del vehículo, este módulo suele estar gestionado por la ECU del Battery Management System (BMS).

2.4.3.2 Importancia en la carga segura

Este sistema de comunicación es el que permite, por ejemplo:

- Que una estación interrumpa la alimentación si el vehículo no está correctamente conectado.
- Ajustar automáticamente la potencia de carga según la disponibilidad energética o el estado térmico de la batería.
- Detener la carga ante fallos de conexión o sobrecalentamiento.

Este mecanismo también es la base sobre la cual se implementan sistemas más avanzados, como el “Plug & Charge” (ISO 15118), que permite la autenticación y el pago automático del servicio de carga a través del mismo canal de comunicación.

2.4.4. Infraestructura de carga de EVs en Uruguay

Si bien la estandarización de los conectores a nivel internacional ha permitido una mayor interoperabilidad entre vehículos eléctricos y estaciones de carga, la infraestructura de carga varía significativamente según la región. En Uruguay, el despliegue de la red de carga ha sido liderado principalmente por UTE, que ha implementado estaciones compatibles con los estándares más utilizados a nivel mundial. La distribución de estos cargadores y la compatibilidad con los vehículos presentes en el país juegan un rol clave en la adopción de la electromovilidad, influyendo en la comodidad y accesibilidad para los usuarios.

En la página web oficial de UTE se encuentra un glosario de información sobre los cargadores que ellos han instalado y mantienen a día de hoy. Allí se puede encontrar información sobre los modos de carga que ofrecen en distintos puntos del país, un mapa interactivo en el que buscar los cargadores o SAVEs disponibles (Sistemas de Alimentación de Vehículos Eléctricos) y una guía de pasos para utilizar uno de estos sistemas públicos [9].



Figura 2.11: Ejemplo de SAVE instalado [9]

El SAVE es el dispositivo intermediario entre la red eléctrica y el vehículo eléctrico y cumple el rol de gestionar la carga del mismo, priorizando la seguridad de los usuarios. En estos es posible encontrar algunos enchufes tipo Schuko o, en su mayoría, encontrar el conector Tipo 2 o Mennekes (y su versión para carga continua, CCS2) vistos en la sección anterior. Sin embargo, es posible encontrar algún conector de tipo GB/T esporádicamente

en el mapa. Además, es posible observar la cantidad de cargadores disponibles en cada punto de carga y las potencias de cada uno.

A fines de noviembre de 2024 se instalan en Uruguay los primeros cargadores ultrarrápidos, gracias a los cargadores de alta velocidad de Huawei, los cuales cuentan con una potencia de hasta $600kW$. Se le ha adjudicado el nombre de "electrolinera" y posee la capacidad de cargar hasta 12 vehículos al mismo tiempo. Hasta la fecha de publicación de este artículo [37], la red de carga uruguaya cuenta con 328 puntos de carga en más de 80 localidades del país y de esta cifra, casi el 40% de ellos permiten la carga rápida. Para promover la adopción de la movilidad eléctrica en el país, UTE se plantea el desafío de instalar un punto de carga cada 50 kilómetros.

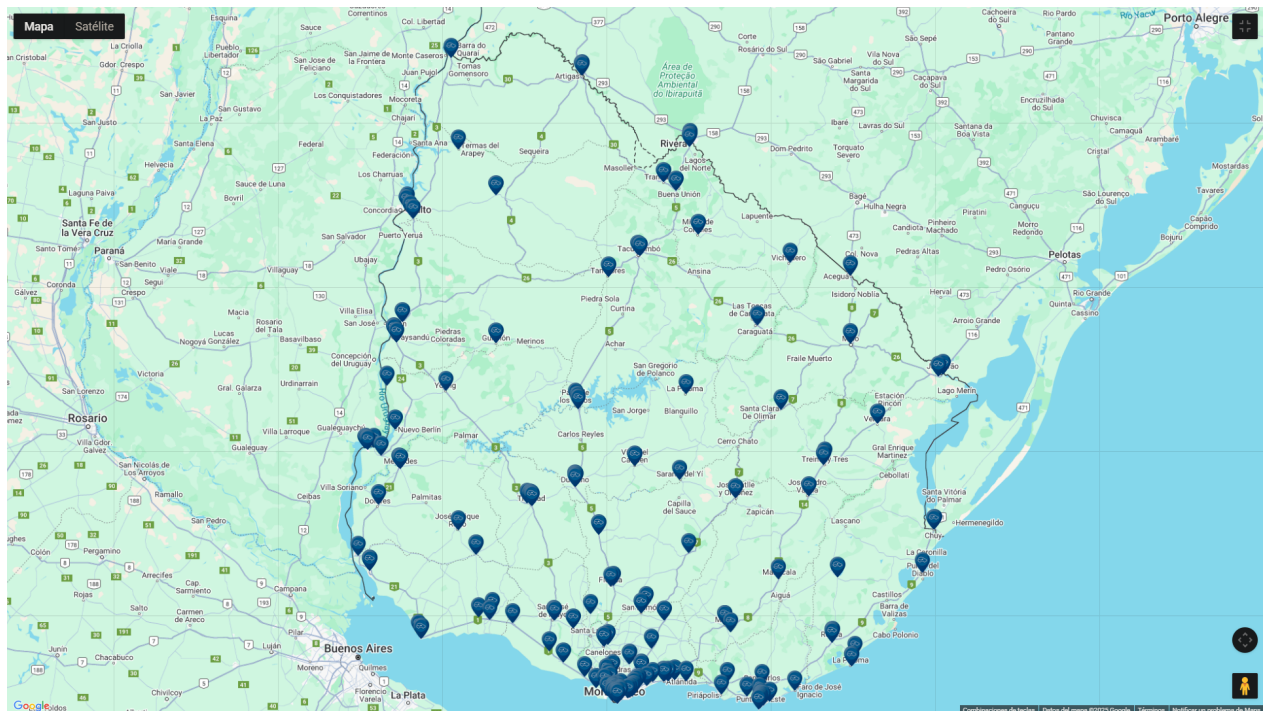


Figura 2.12: Mapa de puntos de carga en Uruguay [9]



Figura 2.13: Ejemplo de punto de carga

2.4.5. Frenado regenerativo

Además de las estaciones de carga, los vehículos eléctricos cuentan con sistemas que optimizan el uso de la energía almacenada en la batería. Uno de los más relevantes es el frenado regenerativo, que permite recuperar parte de la energía cinética del vehículo y convertirla en electricidad, mejorando la autonomía sin necesidad de recurrir a la red de carga externa. Este mecanismo es especialmente útil en entornos urbanos y en regiones con infraestructuras en desarrollo, donde cada oportunidad de optimizar la energía se traduce en una mayor eficiencia del sistema.

2.4.5.1 Funcionamiento

En condiciones normales de conducción, el motor eléctrico del EV transforma la energía eléctrica almacenada en la batería en energía mecánica para mover el vehículo. Durante la desaceleración o al pisar los frenos, el sistema de frenado regenerativo ejecuta el proceso inverso, donde el motor eléctrico actúa como generador y convierte la energía cinética en eléctrica. De esta manera se aumenta la eficiencia del vehículo [38]. Esta característica es extremadamente útil en ambientes urbanos debido al mayor uso del freno, en contraposición a entornos más rurales o en rutas.

2.4.5.2 Beneficios

Al recuperar energía con la desaceleración, se reduce la demanda total de energía, aumentando la autonomía del vehículo y la eficiencia del mismo, reconvirtiendo la energía en vez de perderla en forma de calor. Sumado a esto, está estudiado que el uso de esta característica reduce el desgaste de los frenos mecánicos, ya que el vehículo aumenta su desaceleración con la transformación de la energía. Sin embargo, el frenado regenerativo no sustituye por completo el sistema de frenado mecánico; estos siguen siendo necesarios en casos de frenados de emergencia o cuando se necesita una detención completa.

3. Primera etapa: Obtención de datos de vehículos

Una vez realizada la introducción a las capacidades mecánicas y energéticas de los vehículos eléctricos, fue necesario comprender las capacidades de comunicación que el vehículo posee. Como todo vehículo moderno, los EVs cuentan con múltiples sistemas electrónicos interconectados que permiten monitorear y controlar distintos parámetros operativos en tiempo real.

Desde hace décadas, la evolución de los sistemas de diagnóstico y comunicación en la industria automotriz ha permitido el desarrollo de estándares como el **OBD (*On-Board Diagnostics*)**, protocolos de comunicación como **CAN (*Controller Area Network*)**, y sistemas inteligentes de gestión como el **BMS (*Battery Management System*)**. Estas tecnologías han sido fundamentales para la integración de sensores, la gestión de la batería y la optimización del rendimiento del vehículo.

Para entender mejor la relación entre estos sistemas, a continuación, se presenta una línea temporal que resume su evolución y adopción en la industria automotriz.

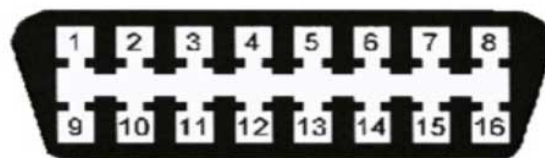
3.1. Evolución de los sistemas de comunicación y diagnóstico en vehículos eléctricos

A lo largo de la historia, los sistemas electrónicos en los vehículos han evolucionado para proporcionar mayor control, diagnóstico y eficiencia. La siguiente cronología destaca los hitos clave en el desarrollo de estos sistemas:

- **1957 - Surgimiento del primer ECU (*Electronic Control Unit*):** Se atribuye a un sistema denominado "*Electrojector system*" desarrollado por *Bendix Corporation*. Fue diseñado en un intento de un sistema de inyección electrónica de combustible. Esta ECU procesaba las señales de distintos sensores para gestionar dicha inyección [39].
- **1983-1986 - Bosch desarrolla CAN:** Con el objetivo en mente de reducir la cantidad de pares de cables torneados y reducir los costos de ensamblaje, Bosch comienza con el desarrollo de un protocolo de comunicación utilizando buses seriales para su enrutamiento y, en febrero de 1986, la Sociedad De Ingenieros Automotrices (en inglés *SAE: Society of Automotive Engineers*) aprueba este estándar [40].

- **1988 - Creación de OBD-I:** Según [41], el grupo CARB (*California Air Resources Board*) lanzó OBD-I, un sistema diseñado para el monitoreo de emisiones. Este sistema de diagnósticos a bordo fue pionero en el mundo del diagnóstico vehicular, aunque faltaba de estandarización de códigos de errores entre los distintos fabricantes.
- **1991 - Implementación de CAN por primera vez:** Mercedes-Benz desarrolla el primer vehículo en implementar el protocolo CAN, utilizando ECUs para el control del motor [42].
- **1994 - ISO 11898-1 estandariza CAN:** La normativa ISO 11898, dividida en cuatro documentos, marca como estándar el protocolo de comunicación CAN para vehículos. El primero de estos documentos ISO 11898-1 incluye la descripción de las capas de enlace de datos y capa física [43].
- **1996 - OBD-II se convierte en estándar para vehículos ligeros:** En EE.UU., la Agencia de Protección Ambiental (EPA) y la SAE estandarizan OBD-II, con un puerto de diagnóstico universal (SAE J1962 3.1) y códigos de error estandarizados (DTC, *Diagnostic Trouble Codes*). Además, en esta versión de OBD se colocó especial énfasis en la seguridad en el vehículo y en el monitoreo del rendimiento. De aquí se desarrollan normas asociadas:
 - SAE J1979: Define los modos de prueba de OBD-II.
 - ISO 15031: Regula la comunicación de diagnóstico

Terminales del Conector OBDII



1 – Sin uso	9 – Sin uso
2 - J1850 Bus positivo	10 - J1850 Bus negativo
3 – Sin uso	11 – Sin uso
4 - Tierra del Vehículo	12 – Sin uso
5 – Tierra de la Señal	13 – Tierra de la señal
6 - CAN High	14 - CAN Low
7 - ISO 9141-2 - Línea K	15 - ISO 9141-2 - Línea L
8 – Sin uso	16 - Batería - positivo

Figura 3.1: Conector SAE J1962, más conocido como conector OBD [10]

- **Década del 2000 - Expansión de CAN en la industria:** La mayoría de los fabricantes adoptan CAN Bus como el protocolo principal para la comunicación entre ECUs en vehículos.
- **2004 - Publicación de ISO 15765-2: Diagnóstico sobre CAN:** Este documento especifica la capa de transporte y capa de red, cumpliendo con los requisitos de sistemas basados en CAN indicados en la norma ISO 11898-1. Posteriormente, se actualiza esta norma para compatibilizar el diagnóstico sobre CAN, según la norma ISO 14229-2 [44].
- **2006 - Introducción ISO 14229-1: *Unified Diagnostics Services* o su sigla, UDS:** Establece los requisitos independientes del enlace de datos para los servicios de diagnóstico, permitiendo que un cliente controle las funciones de diagnóstico en las ECU de los vehículos. La norma completa se compone de 8 documentos, definiendo perfiles según el protocolo de diagnóstico a utilizar [45].
- **2008 - CAN obligatorio para OBD-II en Estados Unidos:** A partir de este año, todos los vehículos vendidos en Estados Unidos están obligados a utilizar CAN como protocolo de comunicación para OBD-II, cumpliendo con las normas ISO 15765 [46].
- **2010 - Desarrollo de BMS avanzados:** Con la creciente adopción de vehículos eléctricos, los BMS evolucionan para monitorear celdas de batería, temperatura y carga, garantizando eficiencia y seguridad.
- **Actualizaciones de normativas ISO:**
 - 2016 - Actualización de ISO 11898-2: Define los estándares eléctricos y de señalización para CAN de alta velocidad (HS-PMA) [47]. En 2024 vuelve a recibir una actualización [48].
 - 2022 - Actualización de ISO 14229-3:2022: Actualización de *UDSonCAN* [49].
 - 2024 - Actualización de ISO 11898-1:2024: Introducción de CAN FD *Light* y CAN XL (*FD, flexible data* y *XL, extra large*) con el objetivo de mejorar la comunicación entre sensores y actuadores [50].

Estos avances han sido fundamentales para la evolución de la movilidad eléctrica, permitiendo a los vehículos modernos recopilar, analizar y compartir datos para optimizar su funcionamiento y facilitar su diagnóstico.

3.2. ¿Qué es CAN?

Controller Area Network (CAN) es un protocolo de comunicación en serie diseñado para permitir la comunicación eficiente y confiable entre múltiples dispositivos electrónicos sin necesidad de una computadora central. Fue desarrollado por Bosch en 1986 y se ha convertido en un estándar en la industria automotriz y otras aplicaciones industriales debido a su robustez, eficiencia y tolerancia a fallos.

El protocolo CAN sigue el modelo OSI (*Open Systems Interconnection*), que divide la comunicación en capas para estructurar y simplificar el intercambio de datos. Aunque CAN opera principalmente en las capas física y de enlace de datos, protocolos como UDS (*Unified Diagnostic Services*) extienden su funcionalidad a niveles superiores.

Según la especificación, el modelo OSI aplicado a CAN y UDS se puede desglosar de la siguiente manera:

Capa OSI	UDS	UDS CAN	en ¿Qué define?
Capa 7 - Aplicación	ISO 14229-1, ISO 27145-3	ISO 14229-3	UDS opera en esta capa para proporcionar servicios de diagnóstico
Capa 6 - Presentación	ISO 27145-2	-	Define codificación específica de datos si es necesario
Capa 5 - Sesión	ISO 14229-2	-	Manejo de sesiones de diagnóstico en UDS
Capa 4 - Transporte	-	ISO 15765-2	ISO-TP permite fragmentación y reensamblado de mensajes
Capa 3 - Red	ISO 27145-4	ISO 15765-2	ISO-TP gestiona la multiplexación de mensajes
Capa 2 - Enlace de datos	-	ISO 11898-1	Define control de acceso al bus y detección de errores
Capa 1 - Física	-	ISO 11898-2, ISO 11898-3	Especifica voltajes, velocidades y topología del bus

Cuadro 3.1: Modelo OSI aplicado a CAN y UDS

En este esquema, CAN se encarga de la transmisión confiable de datos en las capas más bajas, mientras que UDS opera en capas superiores para ofrecer servicios avanzados de diagnóstico y comunicación.

3.2.1. ¿En qué consiste la capa física?

La transmisión de datos en CAN se realiza mediante el bus CAN, compuesto de dos cables trenzados denominados CAN_H (CAN HIGH) y CAN_L (CAN LOW). Los datos se transmiten a través del bus de forma serial utilizando un sistema de tensión diferencial con dos estados posibles, recesivo (equivalente a un 1 lógico) y dominante (equivalente a un 0 lógico) [51]. Cuando ambos cables del bus se encuentran a la misma tensión (2.5V) esto significa que el bus está en estado recesivo, mientras que en estado dominante, el CAN_H se encuentra a mayor voltaje (3.5V) mientras que el CAN_L se encuentra a un menor voltaje (1.5V) [48].

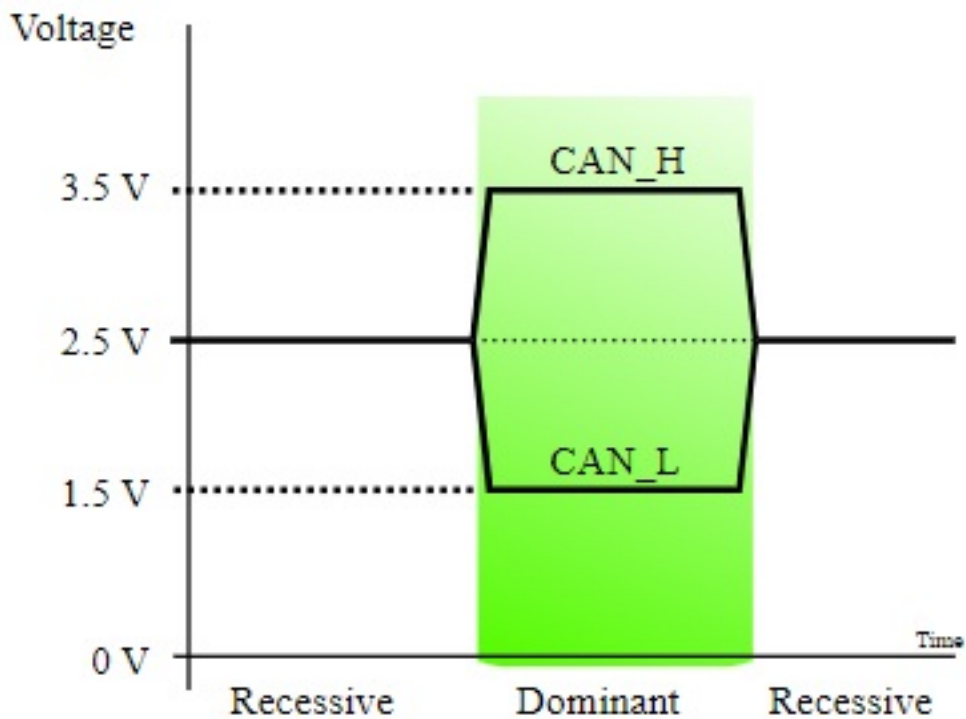


Figura 3.2: Niveles de voltaje de líneas CAN según el estado del bus [11]

3.2.2. ¿Qué es una ECU?

Una unidad de control electrónico (ECU o *electronic control unit* en inglés) es un componente computarizado dentro de los vehículos que controla y administra varios sistemas al recibir información en tiempo real de sensores distribuidos por todo el vehículo. La ECU procesa estos datos utilizando un software integrado que contiene algoritmos adaptados a cada sistema que gestiona. Con base en estos datos procesados, la ECU envía señales de salida a varios actuadores del sistema, que traducen estos comandos en acciones físicas para ajustar los diferentes componentes dentro del vehículo en consecuencia. Normalmente, no hay una sola ECU en un vehículo: hay varias ECU, cada una dedicada a controlar sistemas específicos.

Las ECUs son los nodos de la red CAN, a través de la cuál pueden tanto enviar como recibir datos, y evitar colisiones. Cada ECU tiene tres componentes principales [12]:

- **Microcontrolador:** El microcontrolador recibe información de los sensores y se comunica con los actuadores, además de interpretar los mensajes recibidos por el bus CAN y enviar mensajes a otras ECUs a través del mismo.
- **Controlador CAN:** El controlador CAN se asegura de que toda la comunicación sea consistente con el protocolo CAN, ayudando con la detección de errores, el arbitraje para evitar colisiones, etc.
- **Transceiver CAN:** El *transceiver* CAN conecta al controlador CAN con las conexiones físicas del bus y se encarga de hacer la conversión entre la señal diferencial del bus a los datos que el controlador CAN entiende y envía.

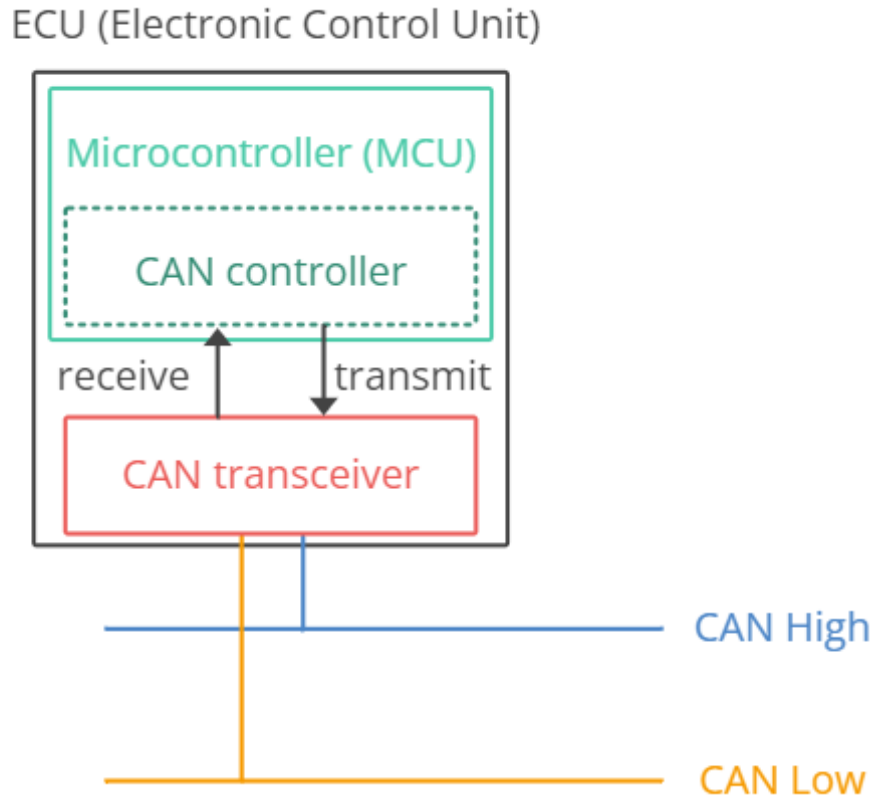


Figura 3.3: Principales componentes de la ECU [12]

3.2.3. ¿Qué es una trama?

CAN es un protocolo basado en mensajes, estos mensajes son conocidos como tramas (o *frames* en inglés). Estas mismas, se definen en la capa de enlace de datos, y están especificadas en la norma ISO 11898-1 [50]. Existen cuatro tipos diferentes de tramas CAN:

- **Trama de datos:** Es la más común de todas, se usa para enviar datos de un nodo a otro, o varios.
- **Trama remota:** Se utiliza para pedir datos de otro nodo CAN enviando una trama que no tiene bytes de datos. En la practica prácticamente nunca se usan ya que para pedir datos se suelen usar protocolos de mayor nivel que utilizan tramas de datos para pedir información
- **Tramas de error:** Se utilizan para indicar errores de comunicación en la red.

- **Tramas de sobrecarga:** Se utilizan en situaciones en que un nodo puede necesitar mas tiempo para procesar información.

En general, en la amplia mayoría de los casos se utilizan solamente tramas de datos y son las que fueron de relevancia para este proyecto. Las tramas de CAN estándar tienen la siguiente estructura:

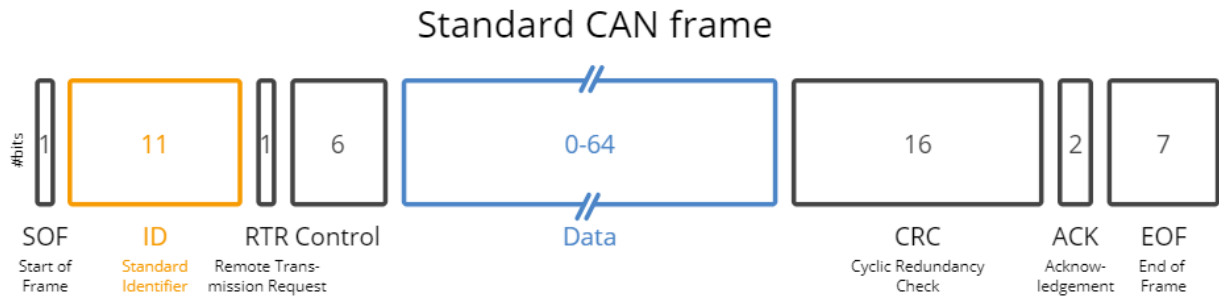


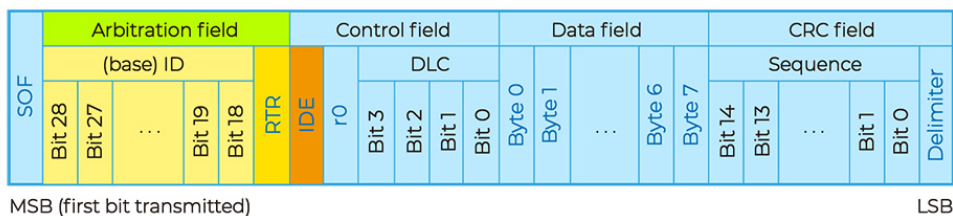
Figura 3.4: Estructura de las tramas de CAN estándar de 11 bits [12]

1. **Start of frame (SOF):** 1 bit, siempre dominante
2. **ID:** 11 bits, identifica al mensaje y a su vez sirve para otorgar prioridad. Si mas de una ECU intenta mandar un mensaje al mismo tiempo, ambas escribirán en el bus hasta que las tramas difieran en un bit. Cuando esto ocurre, el bit dominante siempre toma prioridad por sobre el recesivo y la ECU que envió el bit recesivo deberá esperar. De esta forma. los números mas bajos de ID siempre tienen mayor prioridad
3. **Remote transmission request (RTR):** 1 bit, es dominante en tramas de datos y recesivo en tramas remotas, dándole mayor prioridad a las de datos.
4. **Control:** 6 bits, contiene el bit de extensión de identificador, que es dominante en CAN estándar de 11 bits. También contiene la cantidad de bytes de datos que contiene el mensaje (entre 0 y 8).
5. **Data:** Entre 0 y 8 bytes, contiene la información que se quiere transmitir.
6. **Cyclic redundancy check (CRC):** 16 bits, chequeos de redundancia utilizados para asegurar la integridad de los datos
7. **ACK:** 2 bits, se utiliza para indicar que un nodo ha recibido la información correctamente
8. **End of frame (EOF):** 7 bits, siempre recesivos

3.2.4. Versión alternativa de CAN: 29 bits

La diferencia entre CAN estándar (11 bits) y CAN extendido se encuentra esencialmente en el identificador de la trama. CAN extendido utiliza identificadores de 29 bits, permitiendo una cantidad de identificadores únicos y niveles de prioridad mucho más altos. El CAN extendido es utilizado ampliamente en camiones de gran tamaño y maquinaria pesada en general apoyado de un protocolo especialmente diseñado para CAN extendido que se detalla en la norma SAE J1939. Se aprecian las diferencias entre ambos formatos en la Figura 3.5.

Base CAN data frame format



Extended CAN data frame format

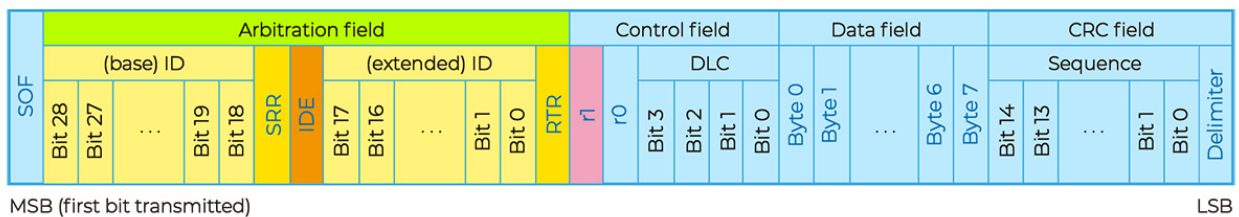


Figura 3.5: Comparación CAN estándar y CAN extendido [13]

3.2.5. Comunicación *multi-frame*

En algunos casos, los 8 *bytes* de datos que una trama de CAN convencional (de 11 bits) tiene disponible pueden no ser suficientes para la información que se desea enviar, en estos casos se deben usar varias tramas de CAN para enviar toda la información. Para esto se definen cuatro tipos de tramas: trama única (*single frame*), primera trama (*first frame*), tramas consecutivas (*consecutive frames*) y tramas de control de flujo (*flow control frames*).

- **Single frame:** El single frame es la trama que se utiliza siempre que la información a enviar no requiera de más de una trama. En el *single frame*, los *bits* 7 a 4 del *byte* 0 son todos 0, y el resto del *byte* se utiliza para indicar el largo del mensaje (de 1 a 7 *bytes*). Este formato reduce el tamaño máximo de información que se puede enviar en

una trama pero permite identificar los mensajes que necesitan mas de una trama de los que no la necesitan.

- **First frame:** Es la primera trama de un mensaje *multi-frame* y sus primeros cuatro *bits* serán 0001. Luego, el resto de los *bits* del *byte* 0 y la totalidad del *byte* 1 se utilizan para indicar el largo del mensaje (de 1 a 4096), dejando solo 6 *bytes* para información dentro de este frame.
- **Consecutive frame:** Las tramas consecutivas tienen un "2" en sus primeros cuatro *bits* (0010), los otros cuatro *bits* del *byte* 0 indican el "número de secuencia"; un número entre 0 y 15 que sirve para que el receptor pueda ordenar los mensajes recibidos. Este número puede reiniciarse a cero si llega a 15 y continuar la comunicación.
- **Flow control frame:** Las tramas de control de flujo son tramas que puede enviar el receptor de los datos para regular cuántas tramas consecutivas quiere recibir en un bloque y cuánto tiempo esperar entre ellas. Esto le permite al receptor, regular la velocidad de entrada de los datos y asegurarse de que tiene tiempo para procesarlos.

	Byte 0	Byte 1	Byte 2	Byte 3	Byte 4	Byte 5	Byte 6	Byte 7	
Single Frame	0	Data Len	Data 0	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6
First Frame	1	Data Len (8-4095)		Data 0	Data 1	Data 2	Data 3	Data 4	Data 5
Consecutive Frame	2	Seq Num	Data 0	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6
Flow Control Frame	3	Flag	Block Size	Separation Time					

Figura 3.6: Tramas definidas para comunicación *multi-frame* [14]

3.2.6. Clasificación de CAN en función de la velocidad de transmisión

Otra forma en la que se pueden diferenciar versiones de CAN es según su velocidad, dependiendo de la aplicación y de las exigencias específicas de cada caso, se definen distintas velocidades de transmisión de los datos:

- **CAN de baja velocidad:** Admite velocidades de hasta 125 Kbps. Es utilizado principalmente en situaciones en las que la velocidad de los datos no es tan crítica, o en situaciones en las que es imperativo que la comunicación funcione a pesar de fallos ya que el CAN de baja velocidad es más robusto ante los fallos.
- **CAN de alta velocidad:** El CAN de alta velocidad es el CAN clásico, es el más común de todos ya que es ideal en situaciones en que se requiere una alta tasa de actualización de la información y en las que conviven varios sistemas al mismo tiempo; es sobre el que se trabajó principalmente en este proyecto. La velocidad máxima es de 1 Mbps.
- **CAN FD:** FD (*Flexible Data-rate*) es la versión más moderna de CAN y admite tasas de transmisión de datos variables. En CAN FD, los campos anteriores y posteriores al campo de datos de la trama se transmiten a 1 Mbps (o menos) como se haría en CAN clásico ya que durante estos campos no necesita sincronización entre todos los nodos para poder mantener las reglas de arbitraje y *acknowledgement*. Pero, en los *bytes* de datos, la velocidad de transmisión se ve delimitada por las características del *transceiver* ya que sólo un nodo estará escribiendo al mismo tiempo. Además de la tasa de transmisión variable, CAN FD también aumenta la cantidad máxima de *bits* de datos de 8 a 64. Esto permite, en conjunto con la velocidad alta de los bytes de datos, velocidades totales de transmisión mucho mayores que las de CAN clásico.

A modo de ejemplo, si se utiliza una velocidad de transmisión 8 veces mayor a la de CAN clásico y se envían 64 bytes en lugar de 8, se estima que la tasa de transmisión final es 6 veces mayor que la original, teniendo en cuenta que las velocidades fuera de los bytes de datos son iguales y que CAN FD tiene algunos *bits* más en estos campos. Si bien en este proyecto no se utilizó CAN FD, a medida que los sistemas electrónicos de los vehículos aumentan y las necesidades de transmisión de información crecen, se espera que se convierta en algo cada vez más común hasta que se convierta en estándar [52].

3.3. ¿Qué es UDS?

Unified Diagnostic Services (UDS) es un protocolo de comunicación de alto nivel utilizado en ECUs principalmente para diagnóstico, reprogramación y actualización de software, testeo de componentes, etc. En vehículos, UDS se implementa sobre CAN, como está especificado en la norma ISO 14229-3 [49] utilizando los *bytes* de datos de las tramas de CAN. Sin embargo, a diferencia de CAN, UDS utiliza un modelo cliente-servidor en el que una herramienta de diagnóstico o escáner (cliente) pide información a una ECU (servidor) y ésta responde adecuadamente.

UDS es un protocolo basado en servicios, un servicio es una acción o comando que el cliente puede solicitar a la ECU. Cada servicio tiene un identificador único de servicio (SID - *Service Identifier*) de 8 *bits* y, en algunos casos, subfunciones o parámetros adicionales. El SID se envía en el *byte* 0 de datos de una trama de datos de CAN, seguida de un identificador de subservicio si es necesario. A continuación, en mensajes de solicitud, se envían parámetros de la solicitud siempre que sea necesario; y en mensajes de respuesta se envían los datos requeridos.

Este proyecto se centra principalmente en el servicio *Read data by Identifier* (SID 0x22) que permite acceder a distintos datos y valores del vehículo. Para esto, se envía en el *byte* 0 el SID 0x22, y en los siguientes 2 se envía un identificador de dato (DID o *data identifier*), que caracteriza el parámetro que se quiere consultar. Luego, el mensaje de respuesta lleva el SID 0x62, seguido del mismo DID y luego el valor del parámetro.

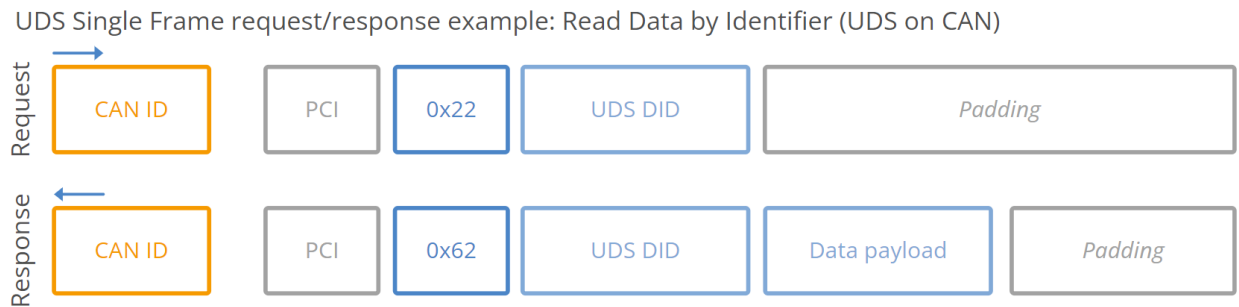


Figura 3.7: Estructura de los mensajes del servicio 0x22 de UDS[15]

3.4. ¿Qué es OBD?

On Board Diagnostics (OBD) es un sistema de diagnóstico a bordo estandarizado en vehículos, que permite la comunicación con las ECUs para el monitoreo y diagnóstico de fallas. Originalmente introducido para monitorear emisiones, su segunda versión expandió el estándar al diagnóstico general del vehículo; estandarizando códigos de diagnóstico de fallas (llamados DTC, *diagnostic trouble codes*) e identificadores de parámetros (PIDs, *parameter identifiers*) para lectura en tiempo real además de fijar a CAN como el protocolo de comunicación base y definir el conector de diagnóstico estándar OBD visto en la figura 3.1.

A día de hoy, OBD se sigue modificando, con actualizaciones a la norma SAE J1979. El estándar SAE J1979-2 (llamado *OBD on UDS*) es una actualización a la norma que establece el uso de UDS como obligatorio para diagnósticos OBD. Esta actualización, introducida en 2021, viene como consecuencia de la alta cantidad de parámetros y complejidades que se deben monitorear en los vehículos modernos. Esto llevó a que ya no existan identificadores de parámetros y DTCs únicos para asignar. Además de la norma SAE J1979-2, se introdujo en 2022 el estándar SAE J1979-3 (*ZEV on UDS*), un complemento a la anterior que establece el diagnóstico a bordo y la utilización de UDS también en vehículos eléctricos, vehículos de cero emisión, y vehículos híbridos enchufables. Este estándar, además de introducir UDS y OBD a los vehículos eléctricos, establece parámetros específicos de los vehículos eléctricos que deben ser medidos así como define DTCs y PIDs característicos de los mismos. Todo esto con el objetivo de optimizar el uso de la energía eléctrica y minimizar las pérdidas de energía mientras esta está almacenada y en su conversión a energía cinética [53].

Estos nuevos estándares demuestran la evolución de los vehículos y su necesidad de diagnóstico y, en particular de los vehículos eléctricos, dado que son un área en constante crecimiento y desarrollo. Sin embargo, estos estándares son muy nuevos y aún no se aplican de manera obligatoria. Particularmente, para el caso de la norma SAE J1979-3, está establecido un período de adopción que resultaría en que todos los vehículos eléctricos lo soporten para 2028. Esto significa que hoy en día, en el rubro de los autos eléctricos no existe estandarización de los PIDs y DTCs específicos a los mismos.[54]

Si bien el nivel de estandarización en vehículos eléctricos deja mucho que desear a día de hoy, y las actualizaciones del estándar aún están en proceso de adopción, la utilización de UDS en ECUs de vehículos ya es común. Muchos vehículos utilizan el protocolo para diagnóstico de parámetros no estandarizados como lo suelen ser los parámetros específicos de vehículos eléctricos.

3.5. Decodificación de datos

Como parte de la investigación realizada, se llevaron a cabo reuniones con diversos proveedores y fabricantes de vehículos eléctricos en el país con el objetivo de obtener acceso a unidades que permitieran el análisis de datos en el contexto de este proyecto. La disponibilidad de vehículos eléctricos resultaba un aspecto fundamental para la ejecución de las pruebas y el desarrollo del estudio.

3.5.1. Acceso a vehículos eléctricos

Para la obtención de datos reales, fue posible contar con la colaboración de **Julio César Lestido S.A.**, empresa importadora y distribuidora oficial de **Volkswagen y Audi** en Uruguay. Gracias a su apoyo, se dispuso del acceso a dos vehículos eléctricos, así como del asesoramiento técnico y el equipamiento necesario para la recopilación y análisis de información. Esta contribución fue clave para el desarrollo de la investigación, permitiendo evaluar en un entorno real las tecnologías y protocolos de comunicación utilizados en estos vehículos.

3.5.2. Proceso de ingeniería inversa

El acceso a los vehículos eléctricos permitió dar inicio al análisis de los sistemas de comunicación internos con el objetivo de extraer y comprender la información relevante para este estudio. Debido a la naturaleza cerrada de muchos de estos sistemas, fue necesario aplicar técnicas de ingeniería inversa, un procedimiento ampliamente utilizado en el ámbito del diagnóstico automotriz y la ciberseguridad vehicular.

La ingeniería inversa en este contexto consiste en interpretar los mensajes transmitidos dentro de la red de comunicación del vehículo, identificando estructuras de datos, protocolos utilizados y parámetros relevantes. Este proceso se basa en la captura y análisis de datos provenientes de buses de comunicación CAN mediante herramientas especializadas.

En el ámbito de la ciberseguridad automotriz, este tipo de análisis se conoce como *hacking ético*, ya que permite acceder a información interna del sistema con fines de estudio y diagnóstico, respetando los principios de seguridad y privacidad de los datos. A continuación, se describirá el procedimiento teórico aplicado para la lectura y decodificación de datos, estableciendo las bases para la formación de la base de datos utilizada en este proyecto.

Este proceso es vital para la identificación de las tramas, para distinguir qué trama contiene qué información. Según las normas visitadas, cada trama posee su propio identificador exclusivo, en la que siempre se envía la misma información. Sin embargo, es raro encontrar una lista de identificadores comunes, ya que es a elección de los fabricantes decidir si un parámetro se envía con el ID 1 o el ID 57.

3.5.3. *Hardware y software utilizado*

Para llevar a cabo este proceso de ingeniería inversa y lograr comprender qué parámetro se encuentra en qué trama y cómo este dato está codificado, hay una serie de condiciones que deben darse:

1. Primero, es necesario acceder a la comunicación interna del vehículo. Afortunadamente, gracias a la estandarización del conector OBD, se logra encontrar fácilmente el acceso a este bus. Generalmente los vehículos suelen tener un conector OBD hembra accesible desde el interior del vehículo, como se ve en la Figura 3.8.



Figura 3.8: Conector OBD en vehículo [16]

2. Supongamos que se busca medir la velocidad, aquí existen dos alternativas para obtener el dato: se hace variar la velocidad conduciendo el vehículo para luego identificar la trama o se simula el envío de esta trama. En ambos casos es necesario identificar entre todas las tramas enviadas cuál es la que contiene el parámetro de velocidad y, sobre todo, cómo está codificado el dato para obtener el valor "real" del parámetro. En el caso de la velocidad es sencillo variarlo de manera externa, pero el problema se encuentra cuando buscamos decodificar parámetros como voltaje y corriente que son claves para el análisis posterior. Además, el dato a buscar puede que no se envíe constantemente a través del bus de datos.

El proceso de ingeniería inversa para la identificación y decodificación de datos en la red de comunicación del vehículo requirió no solo el acceso físico a los buses de datos, sino también herramientas y conocimientos específicos sobre el funcionamiento interno de los sistemas vehiculares. En este sentido, la empresa **Smartway** brindó un apoyo fundamental, proporcionando tanto equipos especializados como recursos digitales que permitieron el análisis de la información transmitida en la red de comunicación del vehículo.

Entre los materiales físicos proporcionados, se incluyeron herramientas de captura y análisis de datos vehiculares, mientras que en el ámbito digital, se tuvo acceso a documentación técnica, *softwares* que la empresa utiliza para la interpretación de mensajería CAN y bases de datos que permitieron agilizar el proceso de identificación de parámetros clave. Gracias a esta colaboración, fue posible optimizar la metodología de decodificación y mejorar la precisión en la interpretación de las tramas de datos, lo que resultó esencial para el desarrollo de este proyecto.

A continuación, se detallan los dispositivos y programas utilizados en el proceso, junto con la metodología aplicada para la recopilación y análisis de los datos obtenidos 3.9.

Desde el vehículo, se precisa un conector en Y para conectar tanto la herramienta de diagnóstico de Volkswagen (ODIS) como la herramienta de decodificación de Smartway (*ValueCAN* y *VehicleSpy*). Por un lado, se obtiene control sobre el vehículo y la comunicación, mientras que por el otro se inspeccionan las tramas en tránsito para su decodificación.

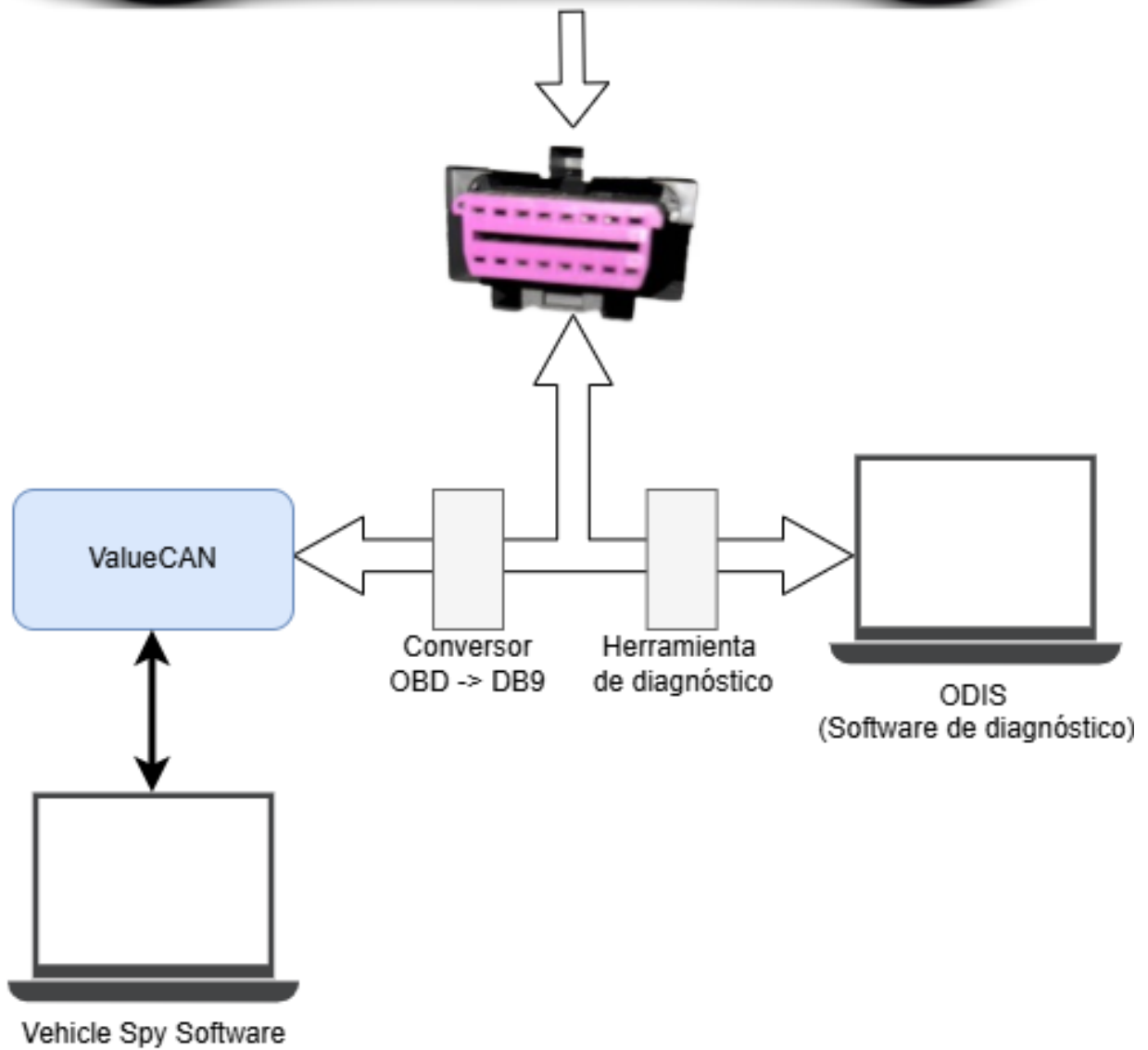


Figura 3.9: Diagrama del proceso de ingeniería inversa realizado

3.5.3.1 ODIS y herramienta de diagnóstico

The Offboard Diagnostic Information System o su sigla ODIS es un sistema de diagnóstico y programación para vehículos del Grupo Volkswagen. Además de Volkswagen, esto incluye marcas como Audi, Škoda y Seat. Esta herramienta permite que concesionarios y talleres independientes puedan acceder a una cobertura completa de los sistemas electrónicos del vehículo. Dentro de funcionalidades más específicas, incluye realizar diagnósticos rápidos y precisos, leer y borrar códigos de falla, acceder a datos en tiempo real y ejecutar pruebas guiadas de localización de fallas.

Además del diagnóstico, ODIS cuenta con capacidades avanzadas de codificación y programación, lo que permite la configuración de módulos electrónicos, la reprogramación de unidades de control y la actualización de software de los vehículos, garantizando que el mantenimiento se realice conforme a los estándares del fabricante. El software se mantiene actualizado de manera periódica para asegurar compatibilidad con los modelos más recientes y proporcionar las herramientas más avanzadas para el análisis y solución de problemas en el sistema eléctrico y electrónico de los vehículos [55].

De todas estas funcionalidades, la que más provecho se le sacó es la opción de leer parámetros particulares dentro de una ECU. Esta función es esencial ya que permite dos procedimientos clave para la decodificación:

1. Primero, contrastar los valores vistos en el *ODIS* contra las tramas vistas en el *VehicleSpy*. Si se observa el parámetro objetivo variar en el *ODIS*, se puede observar al mismo tiempo el inspector de tramas y revisar cuál o cuáles se vieron modificadas.
2. Segundo y más crucial, se utilizará como confirmación de la correcta decodificación de parámetros (se explica en la siguiente sección con las funciones del *VehicleSpy*).

3.5.3.2 ValueCAN y VehicleSpy - Hardware y software de decodificación

Para la decodificación de las tramas de comunicación dentro del vehículo, fue necesario utilizar herramientas especializadas en la captura y análisis de datos en redes CAN. En este contexto, se utilizó el ValueCAN, un dispositivo de interfaz de comunicación desarrollado por *Intrepid Control Systems*, diseñado para interactuar con redes automotrices y capturar información en tiempo real.



Figura 3.10: Herramienta *ValueCAN* [17]

ValueCAN actúa como un puente entre la red de comunicación del vehículo y una computadora, permitiendo la visualización y análisis de los mensajes transmitidos. Su funcionalidad es crucial para la identificación de tramas relevantes dentro del flujo de datos, ya que permite registrar, filtrar y transmitir información específica desde y hacia la ECU.

Para el análisis y procesamiento de la información obtenida a través de *ValueCAN*, se utilizó el software especializado ***VehicleSpy***. Este programa permite realizar múltiples funciones clave en la decodificación de datos, tales como:

- **Captura y almacenamiento de tramas:** *VehicleSpy* permite registrar la información transmitida en el bus de datos del vehículo, almacenando las tramas en una base de datos para su análisis posterior. Además permite guardar una "grabación" de toda la información que se observa cuando se está analizando la comunicación en vivo.
- **Visualización y filtrado de datos:** A través de su interfaz gráfica, el software permite aislar tramas específicas, facilitando la identificación de aquellas que contienen los parámetros de interés.
- **Carga de bases de datos de mensajes:** *VehicleSpy* admite la importación de archivos con estructuras de datos predefinidas, lo que permite interpretar de manera más eficiente la información transmitida.
- **Decodificación de parámetros:** Mediante el análisis en tiempo real de los datos observados en ODIS y en las tramas capturadas, se pueden correlacionar valores específicos con mensajes CAN, facilitando la comprensión de la estructura de la comunicación interna del vehículo.

En la Figura 3.11, se muestra una captura de pantalla de una de estas grabaciones en la que se observan los parámetros siendo decodificados (según la base de datos cargada). En la imagen se observa que el vehículo está siendo cargado con corriente alterna (AC), que la corriente entrante (dividida en 3 fases) vale $16A$ en la primera fase y nula en las restantes, y la eficiencia energética de un 97% .

Filter	Count	Time (abs/rel)	Tx	Er	Description	Arbid/Header	Len	DataBytes	Network	Node	ChangeCnt	Timestamp
	10	1.999931 s			HS CAN \$700 - RQST KeepAlive	700	8	02 3E 80 55 55 55 55 55	HS CAN		0	2024/06/20 12:10:26:1767
	553	553.166 ms			HS CAN \$744 - RQST Cargador	744	8	03 22 1D A0 55 55 55 55	HS CAN		0	2024/06/20 12:10:25:810878
	789	789.490 ms			HS CAN \$744 - RQST Power efficiency	744	8	03 22 15 D6 55 55 55 55	HS CAN		0	2024/06/20 12:10:26:290556
	789	789.630 ms			HS CAN \$744 - RQST AC input current	744	8	03 22 41 FB 55 55 55 55	HS CAN		0	2024/06/20 12:10:26:211288
	793	793.511 ms			HS CAN \$744 - RQST AC input voltage	744	8	03 22 41 FC 55 55 55 55	HS CAN		0	2024/06/20 12:10:26:234969
	553	553.000 ms			HS CAN \$7AE - Rx Cargador	7AE	5	04 62 1D A0 01	HS CAN		0	
					Cargador conectado + tipo							
					= Cargando AC [1]							
	789	789.000 ms			HS CAN \$7AE - Rx Power efficiency	7AE	5	04 62 15 D6 DC	HS CAN		0	
					Power efficiency							
					= 97.0 % [DC]							
	1863	10.000 ms			HS CAN \$7AE - Rx AC input current	7AE	7	06 62 41 FB A0 00 00	HS CAN		0	
					Modulo 1							
					= 16.0 A [A0]							
					Modulo 2							
					= 0.0 A [0]							
					Modulo 3							
					= 0.0 A [0]							

Figura 3.11: Dispositivo ValueCAN y software VehicleSpy en funcionamiento

El uso combinado de ValueCAN y VehicleSpy permitió obtener una visión detallada de la comunicación interna del vehículo y establecer una metodología efectiva para la decodificación de los parámetros de interés. En las siguientes secciones, se detallan los resultados obtenidos de dos ejemplares eléctricos, puestos a disposición por el equipo de Julio César Lestido.

3.5.4. Volkswagen e UP! (auto eléctrico)

El Volkswagen e-up! es un vehículo 100% eléctrico lanzado comercialmente por primera vez en 2013 como una variante del modelo up!. En su versión más reciente, el e-up! cuenta con una batería de $36,8$ kWh de capacidad útil y un voltaje de trabajo nominal de 311 V, lo cual le otorga una autonomía estimada cercana a los 260 km.

Se realizaron varias visitas a la sede central de Julio César Lestido en la calle Cerro Largo, se analizaron siete ECUs diferentes, diversas estructuras de mensajes y métodos de decodificación. Antes de presentar la información, parece interesante mencionar el siguiente diagrama 3.12. En el proceso de descubrimiento de las funcionalidades del sistema *ODIS* se encuentra un mapa más complejo que el presentado, que indica la presencia de una unidad de control intermedia de la comunicación y esta realiza los pedidos de información al resto de las unidades de control. En cada grupo se encuentran entre tres y quince ECUs.

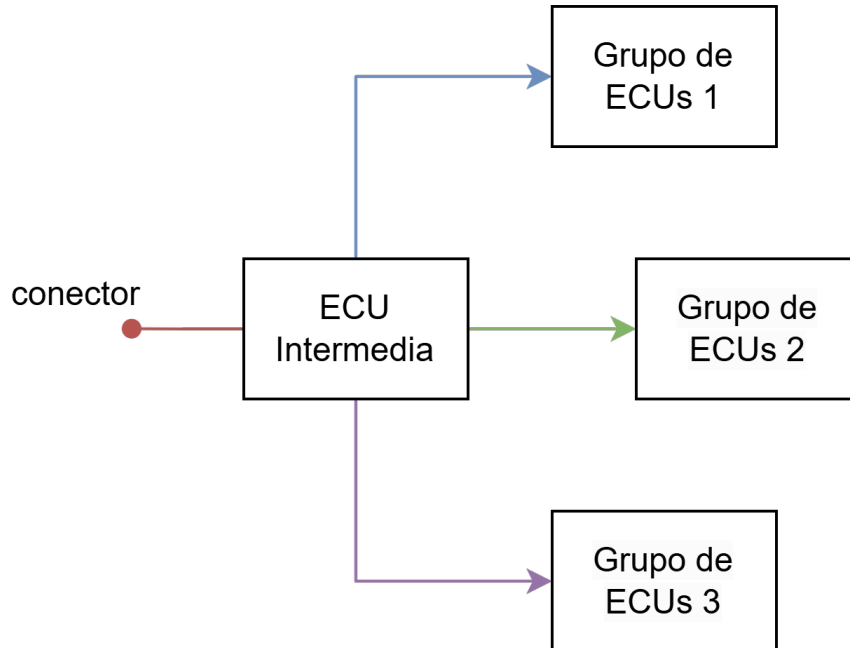


Figura 3.12: Mapa representativo de ECUs en Volkswagen e UP

Respecto a la información del vehículo, se encontrarán con tablas con los parámetros obtenidos y su decodificación en el anexo A.1. Aquí se destacarán aspectos claves al proceso de ingeniería inversa y para su posterior análisis.

Primero, destacar que la comunicación CAN se dio entre los pines 6 y 14 del conector OBD, ubicación clásica para este bus, a una velocidad de transmisión de $500kbps$. Sin esta información, todo el proceso de ingeniería inversa sería prácticamente imposible de realizar.

Otro comentario clave de los resultados es que se obtuvieron parámetros que no parecerían (a simple vista) para la estimación del estado de salud, motivante del proyecto. Esto es correcto, algunos de estos parámetros, como la tensión en bornes +15, +30 y +50 (de ahora en más, parámetros auxiliares), son vitales para el funcionamiento del dispositivo *IoT* a instalar. En la siguiente sección se comentará más al respecto.



Figura 3.13: Foto analizando el Volkswagen e UP

En el proceso de ingeniería inversa se definen dos tareas:

- La primera consiste en identificar la ECU que contiene el parámetro a encontrar y los identificadores correspondientes (tanto de la ECU como del parámetro).
- La segunda, una vez encontrada la trama en la que se envía el valor del parámetro, identificar que *bytes* o qué *words* representan al parámetro, en especial cuando en el mismo mensaje se envía más de un dato.

Para la primera tarea, la tarea se ve facilitada ya que el ODIS contiene parte de esta información que se busca. Sin embargo, para aprovechar al máximo el tiempo de cada visita al vehículo, se encuentra una página web de un foro alemán con información de este vehículo identificada por ECU [56]. Esto permitió paralelizar el trabajo, identificando qué parámetros se encuentran en cada ECU.

Haber resuelto la primera tarea de manera sencilla permitió dedicar más esfuerzos en la segunda tarea de decodificar los parámetros. Dado que el sistema ODIS no permitía acceder directamente a la decodificación el proceso fue el siguiente. Se utilizará de ejemplo la variable "Datos ambientales 1" que se encuentra en la tabla A.5 del anexo A.1:

Se sabía que el kilometraje se encontraba en esta trama, para decodificarla se utilizó el modo "Run with Transmit" del *VehicleSpy* que permite enviar tramas modificadas al bus de datos. Al enviar datos al bus, estos serán captados por el ODIS y se observa cómo se altera el valor decodificado.

- Borrador de averías sin confirmar: 0
- Kilometraje: 35.235 km
- *VehicleSpy*: 07 62 02 BD 00 89 A3 55

Se reconoce el largo de la información en la trama (07), el 62 como identificador de respuesta del servicio 22 y, por ende, las siguientes dos *words* deben representar el identificador del parámetro.

Después de tanto análisis, ya se había descubierto que este vehículo utilizaba la *word* 55 como relleno de mensajes (dada la estructura de CAN de 11 bits presentada es necesario que el contenido tenga ocho *words*). Entonces quedan tres *words* a observar. Se envía a través del *VehicleSpy* la trama utilizando el servicio 22 y se observa en ODIS la siguiente trama que responde con el byte 62:

- *VehicleSpy*: 07 62 02 BD 10 89 A3 55
- Borrador de averías sin confirmar: 16
- Kilometraje: 35.235 km

Con este ejemplo y otros similares modificando esta *word* se concluye que esta *word* contiene en formato hexadecimal al contador de averías. Ahora, se repite el proceso modificando la siguiente *word*:

Envío 1:

- *VehicleSpy*: 07 62 02 BD 00 8A A3 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 35.491 km

Envío 2:

- *VehicleSpy*: 07 62 02 BD 00 88 A3 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 34.979 km

La diferencia presentada entre modificar 89 a 88 y 89 a 8A es la misma en el valor en kilómetros:

$$35,491 - 35,235 = 35,235 - 34,979 = 256$$

Variando el bit menos significativo de este *byte* modifica en 256 kilómetros, para terminar de confirmar la decodificación se enviaron un par más de tramas:

Envío 3:

- *VehicleSpy*: 07 62 02 BD 10 89 A2 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 35.234 km

Envío 4:

- *VehicleSpy*: 07 62 02 BD 10 88 A2 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 34.978 km

Envío 5:

- *VehicleSpy*: 07 62 02 BD 10 8A A2 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 35.490 km

Envío 6:

- *VehicleSpy*: 07 62 02 BD 10 89 A4 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 35.236 km

Envío 7:

- *VehicleSpy*: 07 62 02 BD 10 88 A4 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 34.980 km

Envío 8:

- *VehicleSpy*: 07 62 02 BD 10 8A A4 55
- Borrador de averías sin confirmar: 0
- Kilometraje: 35.492 km

Estos resultados permiten confirmar que la decodificación del kilometraje se representa con dos *bytes* consecutivos. Este resultado se puede observar en la tabla A.2 del anexo A.1.

3.5.5. Volkswagen e Delivery (camión eléctrico)

Con grandes avances realizados en el primer vehículo eléctrico, se decide ampliar el alcance del proyecto incorporando un segundo modelo de vehículo eléctrico, permitiendo así un análisis comparativo entre diferentes tipos de arquitectura vehicular, en este caso, un automóvil y un camión eléctrico. Al igual que el Volkswagen e UP, el Volkswagen e Delivery fue puesto a disposición por el equipo de Julio César Lestido en conjunto con una versión diferente del sistema ODIS para realizar el proceso de ingeniería inversa.

El Volkswagen e-Delivery es un camión eléctrico desarrollado por Volkswagen Caminhões e Ônibus en Brasil, diseñado específicamente para el transporte urbano de carga con cero emisiones. Fue presentado en 2017 y lanzado al mercado en 2021 como parte de la estrategia de electrificación de la línea de vehículos comerciales. En el caso particular del vehículo analizado en este proyecto, se trató de una versión equipada con seis packs de batería de iones de litio, cada uno con una capacidad de 35 kWh, alcanzando así una capacidad total de 210 kWh y operando a un voltaje nominal de 650 V. Esta configuración le permite una autonomía estimada de hasta 250 km en entornos urbanos.

Se presentaron varias diferencias en comparación al proceso de ingeniería inversa realizado en el vehículo anterior:

- Primero, al conectar por primera vez el software *VehicleSpy* en el puerto OBD del camión se encuentran mensajes con información del vehículo. La norma J1939, escrita para *heavy-weight vehicles*, define una serie de mensajes que se envían tipo "broadcast", para que todas las ECU del vehículo la reciban. Esta información se define a partir de un PGN (*Parameter Group Number*) definidos en la norma. En la página citada [57] se puede encontrar mucha información sobre esta norma y hasta un buscador de PGN (norma SAE J1939-71).

Este problema fue parcialmente resuelto al cargar en el software la base de datos correspondiente a la norma SAE J1939. De los más de 70 mensajes sin decodificar de *broadcast* que se encontraron, solamente 30 de ellos permanecieron sin decodificar. Al revisar los PGN de estas tramas, se descubre que la norma los marca como "Datos de propietario", por ende, no estándar y sería necesario otra forma de verificar qué información contienen. Se referirá en este documento como "Problema 1" a esta situación.

- La segunda diferencia encontrada corresponde en el conector OBD. Si bien, convencionalmente se encuentra una única línea CAN (pines 6 y 14) en el conector, en este

caso se encuentran no dos, sino tres líneas de comunicación CAN, en los pares 3-11 y 12-13 correspondientemente. Encontrar la segunda línea CAN no fue tarea difícil ya que el *ValueCAN* permite decodificar dos líneas simultáneamente. En la imagen de la herramienta 3.10 se observa claramente las dos líneas (CAN 1 y CAN 2).

Previo al descubrimiento de la tercera línea CAN, se intentó el envío de tramas utilizando el servicio 0x22 para observar los parámetros en acción. Sin embargo, no se reciben nuevos mensajes en el *VehicleSpy*. Se referirá en este documento como "Problema 2" a esta situación.

3.5.5.1 Problema 1: Mensajes *broadcast* sin decodificar

Inicialmente, al conectar el software *VehicleSpy* al puerto OBD del camión, se identificaron múltiples mensajes de tipo *broadcast*, definidos por la norma SAE J1939. Estos mensajes contienen información estandarizada transmitida a todas las ECUs del vehículo y están organizados mediante PGN (*Parameter Group Number*) según la especificación de la norma.

Para abordar esta situación, se cargó en *VehicleSpy* la base de datos correspondiente a la norma SAE J1939, lo que permitió decodificar gran parte de los mensajes observados. Sin embargo, aproximadamente 30 tramas permanecieron sin decodificar, y al analizar sus PGN, se descubrió que la norma los clasificaba como "Datos de propietario", es decir, mensajes específicos del fabricante y no documentados en el estándar. Esto representó un desafío, ya que la decodificación de estos datos requería un método adicional para identificar su contenido.

Ante esta dificultad, se decidió utilizar el sistema ODIS con el objetivo de enviar mensajes específicos a la ECU y observar si se generaban nuevas tramas en la red CAN del vehículo. No obstante, al realizar este procedimiento, no se identificaron variaciones significativas en las tramas observadas en *VehicleSpy*, lo que indicaba que la estrategia no era efectiva para descifrar la información contenida en los mensajes no estandarizados.

Dado que no se obtuvieron resultados concluyentes, se optó por un enfoque diferente: modificar manualmente las tramas enviadas y observar los cambios reflejados en el tablero digital del vehículo. A través de este método, se lograron identificar varios parámetros clave, tales como:

- Velocidad del vehículo.

- Consumo energético por kilómetro (kWh/km).
- Nivel de regeneración de energía.
- Kilometraje restante.

Si bien este procedimiento permitió recuperar información relevante, la limitación en la capacidad de solicitar datos específicos mediante ODIS llevó a la necesidad de investigar más a fondo la estructura de comunicación del camión, lo que derivó en la identificación de un segundo problema.

3.5.5.2 Problema 2: Tercera línea CAN

Luego del análisis inicial y tras los resultados obtenidos en la primera etapa, se decidió inspeccionar el mapeo de pines en el conector OBD-II, dado que, por experiencia de Smartway, en algunos casos no todas las líneas CAN se encuentran accesibles en los pines estándar.

Mediante este procedimiento, se descubrió la existencia de una tercera línea CAN, cuyos datos se transmitían a través de los pines 12 y 13. Sin embargo, al conectar *VehicleSpy* a esta línea, se detectaban datos aparentemente erróneos o "basura", lo que indicaba una posible diferencia en la velocidad de transmisión en comparación con los dos buses CAN previamente identificados.

Para resolver esta situación, se utilizó la función de calibración automática de *VehicleSpy*, que permite ajustar el hardware de captura (*ValueCAN*) a la velocidad de transmisión del bus de datos. A través de este proceso, se determinó que la tercera línea no operaba a $250kbps$ como los dos primeros buses CAN, sino que utilizaba una velocidad de $500kbps$. Una vez ajustada la configuración, fue posible visualizar nueva información en esta línea de comunicación.

Con la configuración corregida, se volvió a utilizar ODIS para intentar obtener parámetros adicionales a partir del nuevo bus CAN descubierto. Sin embargo, tras varios intentos, se comprobó que ODIS no proporcionaba acceso a los datos clave requeridos para este estudio, tales como:

- Voltaje de la batería.
- Corriente suministrada o consumida.
- Temperatura de la batería.

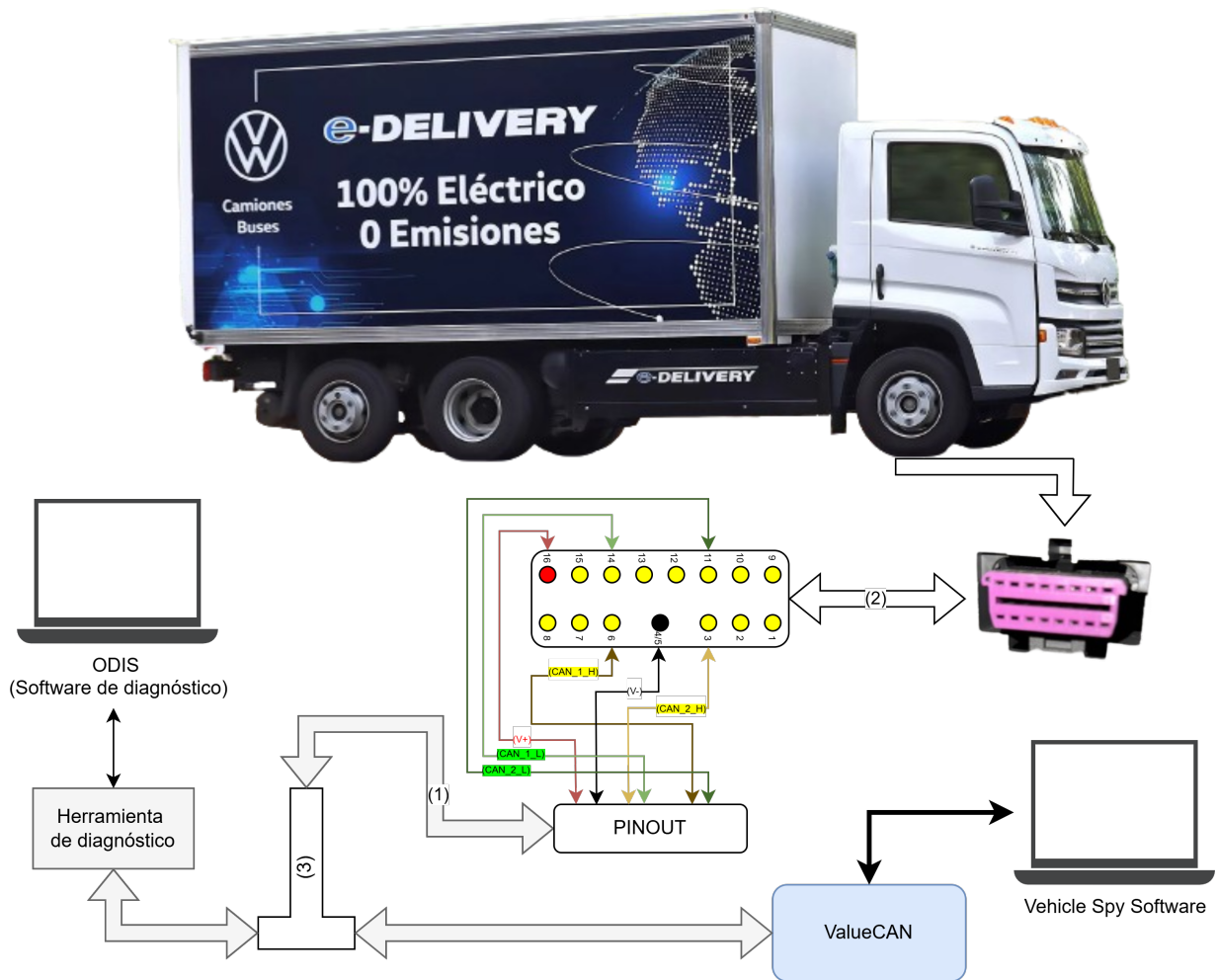
El único parámetro relevante disponible en ODIS fue el *State of Charge* (SoC), pero no se obtuvo acceso a los demás valores necesarios para evaluar el estado de la batería del camión.

Ante la imposibilidad de recuperar los datos esenciales para el análisis, se decidió descartar temporalmente la decodificación del Volkswagen e-Delivery y centrar los esfuerzos en el acondicionamiento del dispositivo de recopilación de datos en el Volkswagen UP! eléctrico, con el fin de continuar con la implementación y pruebas del sistema en un entorno controlado.

3.5.5.3 Conexiones para ingeniería inversa

Figura 3.14: Diagrama de conexiones para la decodificación de parámetros del camión.

Los conectores numerados de la imagen serán encontrados en el Anexo A.2.



Las diferencias de este diagrama con el diagrama de ingeniería inversa en el Volkswagen e UP son gracias a la presencia de las múltiples líneas CAN presentes en el Volkswagen e

Delivery. Como el *ValueCAN* solamente acepta dos líneas de forma simultánea, la caja con los pines del CAN separados permitió enrutar los buses CAN a gusto. Una vez descartado el primer bus (por falta de información relevante), se utilizan los conectores disponibles y unos cables de cocodrilo para "rearmar" el conector OBD y leer solamente la segunda y tercera línea CAN.

3.5.6. Comparación entre vehículos analizados

El análisis de los dos vehículos eléctricos permitió evidenciar diferencias significativas en la estructura y acceso a la información de sus redes de comunicación. Estas diferencias resaltan una problemática clave en la industria de los vehículos eléctricos: la falta de estandarización en la mensajería CAN y en los identificadores de datos, lo que dificulta la interoperabilidad y el diagnóstico universal.

Uno de los principales hallazgos de este estudio es que, a diferencia de los vehículos de combustión interna, los vehículos eléctricos presentan una mayor heterogeneidad en la forma en la que transmiten sus parámetros internos. En vehículos a combustión, los estándares OBD-II y SAE J1979 han logrado una normalización considerable en la lectura de parámetros como revoluciones por minuto, temperatura del motor y consumo de combustible. Sin embargo, en el caso de los vehículos eléctricos, no existe una uniformidad en la definición de parámetros críticos como el voltaje de la batería, la corriente de carga/descarga o la temperatura del sistema de almacenamiento energético.

En el caso del Volkswagen e UP!, la decodificación de parámetros resultó más accesible debido a que el vehículo utiliza una única línea CAN de 500 kbps con un esquema de mensajería compatible con herramientas de diagnóstico como ODIS. Esto permitió la identificación y validación de múltiples variables a través del contraste entre ODIS y VehicleSpy, facilitando la extracción de datos útiles para su posterior análisis e implementación en el sistema de recopilación.

Por otro lado, el Volkswagen e Delivery presentó mayores desafíos en términos de acceso a la información. Se identificaron tres buses CAN diferentes, con velocidades de transmisión variables, y una gran cantidad de tramas clasificadas como datos de propietario según la norma SAE J1939. La decodificación se vio limitada debido a la ausencia de identificadores estandarizados para parámetros eléctricos clave, lo que imposibilitó la recuperación de información como el voltaje y la corriente de la batería. Además, el uso de múltiples líneas CAN

obligó a adoptar estrategias más complejas para la captura de datos, requiriendo un mapeo manual de los pines del conector OBD y la calibración de la velocidad de transmisión del bus.

Otro aspecto crítico identificado es la dependencia de los fabricantes en cuanto a la accesibilidad de datos a través de herramientas no invasivas. La estrategia utilizada en este proyecto se basó en un enfoque pasivo de captura de datos, sin realizar modificaciones en la red de comunicación del vehículo ni intervenir en el software de las ECUs. Sin embargo, este enfoque depende en gran medida de la disposición de los fabricantes a exponer la información de sus sistemas mediante herramientas de diagnóstico o bases de datos accesibles. En el caso del e UP!, la estructura de mensajería permitió extraer información clave, mientras que en el e Delivery, la falta de acceso a identificadores específicos impidió obtener datos fundamentales para el análisis de la batería.

En resumen, la comparación entre ambos vehículos refuerza la necesidad de avanzar hacia una mayor estandarización en los datos de vehículos eléctricos, especialmente en lo que respecta a los parámetros de salud de la batería. La falta de una estructura de datos común entre fabricantes dificulta la implementación de sistemas de diagnóstico universales y puede representar una barrera para el mantenimiento y la optimización de estos vehículos a largo plazo.

3.5.7. Resumen

A partir del proceso de ingeniería inversa realizado en los dos vehículos eléctricos, se tomaron diversas decisiones con el objetivo de optimizar la recopilación de datos y maximizar la utilidad del sistema desarrollado en este proyecto.

El análisis del Volkswagen e UP! permitió la identificación y decodificación de una serie de parámetros clave que serán utilizados en el dispositivo de monitoreo. Se logró mapear información esencial, incluyendo velocidad, kilometraje, estado de carga (SoC), voltajes auxiliares y consumo energético. Estos datos serán fundamentales para evaluar el rendimiento del vehículo y alimentar los modelos de estimación de estado de salud de la batería.

Por otro lado, en el Volkswagen e Delivery, a pesar de los esfuerzos realizados, no fue posible recuperar información relevante sobre el sistema de almacenamiento energético debido a la falta de accesibilidad a identificadores específicos en su red de comunicación CAN. Esto

llevó a la decisión de descartar temporalmente el análisis de este vehículo y concentrar los esfuerzos en la integración del dispositivo de recopilación de datos en el e UP!.

Las principales conclusiones obtenidas de esta etapa son:

- La heterogeneidad en los protocolos de comunicación entre vehículos eléctricos dificulta la obtención de información estandarizada.
- La disponibilidad de herramientas de diagnóstico como ODIS puede facilitar la decodificación de parámetros en algunos modelos, pero no garantiza el acceso a datos críticos en todos los casos.
- La exploración de múltiples líneas CAN en vehículos pesados es un desafío técnico adicional que requiere técnicas avanzadas de identificación y calibración.
- La falta de información accesible en algunos vehículos refuerza la necesidad de contar con estrategias de monitoreo de datos que no dependan exclusivamente del acceso a buses de comunicación internos.

Con base en estos hallazgos, la siguiente etapa del proyecto se enfocará en la implementación y acondicionamiento del dispositivo de recopilación de datos en el Volkswagen e UP!, asegurando su integración con la base de datos de parámetros identificados y optimizando su funcionalidad para el monitoreo en tiempo real del estado del vehículo.

3.6. Recopilación de datos

3.6.1. Acondicionamiento del dispositivo

Posterior a completar y refinar la base de datos del Volkswagen UP y dejar en pausa los avances de decodificación del Volkswagen e Delivery, se decide proceder hacia el acondicionamiento del dispositivo a instalar que recopilará y reportará los datos. Se analizaron diferentes dispositivos capaces de analizar el protocolo CAN y reportar datos, como por ejemplo, el AutoPi [58].

El AutoPi es un dispositivo de telecomunicaciones que permite la transferencia de datos CAN u OBD-II hacia USB o WiFi o mediante su API. Este dispositivo está basado en la *Raspberry Pi 3 model A+* y fue utilizado en un proyecto de grado de la Universidad Católica del Uruguay de nombre "Optimización de la Vida Útil de las Baterías en Vehículos Eléctricos a Partir de su Uso" [59].

Afortunadamente, haber contado con el apoyo de Smartway no fue solamente de utilidad para la decodificación de los vehículos, sino que también lo fue para poner a disposición el **dispositivo SG**.



Figura 3.15: Dispositivo SG de Smartway

Una vez consolidada la base de datos con los parámetros decodificados del Volkswagen e UP, fue necesario integrar esta información en el dispositivo SG para su correcta operación. El dispositivo fue configurado con los identificadores de tramas obtenidos en el proceso de ingeniería inversa, de manera que pudiera solicitar y extraer la información relevante del vehículo de forma autónoma. Lamentablemente, este proceso es parte de un procedimiento interno de la empresa, por lo que no es posible detallar información de este proceso.

El funcionamiento del SG se basa en la comunicación con el bus CAN del vehículo, donde consulta los parámetros previamente identificados y los registra de acuerdo con el estado del vehículo. Dependiendo de las condiciones de operación, el dispositivo ajusta la frecuencia de adquisición de datos y la adquisición de qué datos particularmente (no es sensato pedir la velocidad sabiendo que el vehículo se encuentra apagado), optimizando así el uso de los recursos y minimizando la carga sobre la red de comunicación del vehículo.

Una vez obtenidos los datos, el dispositivo se encarga de transmitir la información recopilada a una plataforma web mediante una conexión de red. Este sistema permite el almacenamiento y monitoreo remoto de los datos, facilitando su posterior análisis.

3.6.2. Puesta a prueba del dispositivo

Una vez configurado el dispositivo y previo a su instalación, fue necesario realizar pruebas sobre este dispositivo para verificar su correcto comportamiento. Con el dispositivo instalado en el vehículo se vuelve más difícil el proceso de depuración de errores. Gracias a las capacidades del Vehicle Spy y con el siguiente diagrama de conexiones 3.16, fue posible verificar por completo el funcionamiento del dispositivo:

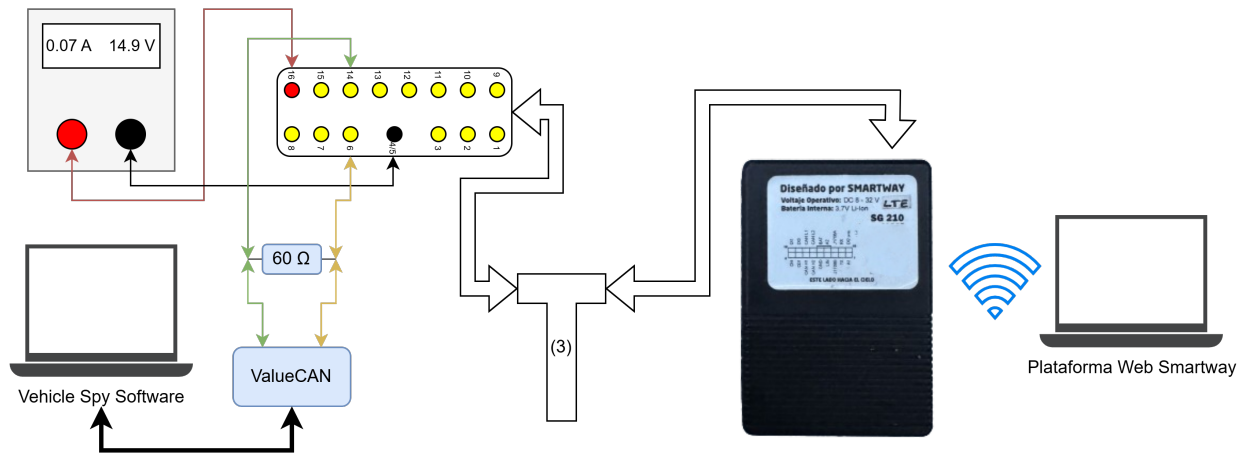


Figura 3.16: Diagrama de sistema de prueba de dispositivo SG de Smartway

El dispositivo maneja cuatro estados:

- **Apagado:** El dispositivo reporta solamente la tensión en el borne +15 cada 5 minutos, además del estado del vehículo que se reporta en todos los estados. Cuando en el borne +15 recibe "ON", el dispositivo pasa al siguiente estado.
- **Moderando:** El vehículo se pone en contacto y comienza a preguntar y reportar por toda la información que contiene en la base de datos cada 30 segundos. Para pasar al siguiente estado, la velocidad debe ser positiva.
- **Conducción:** Reporta la misma información pero cada 10 segundos solamente.
- **Funcionamiento:** Este estado Smartway lo utiliza en aplicaciones agrícolas, en este caso será utilizado como "el vehículo se está recargando" y reportará la información cada 5 segundos. Para detectar este caso, se debe recibir en la variable "Cargador Conectado" un valor distinto de cero (0 = no cargando, 1 = cargando AC y 3 = cargando CC).

Al probar todas estas funcionalidades, se prosiguió con la instalación del dispositivo SG en el auto eléctrico Volkswagen e UP.

3.6.3. Instalación del dispositivo

Se realizaron un par de visitas hasta que finalmente quedó el dispositivo instalado y reportando todos los parámetros correctamente. El principal problema encontrado en el funcionamiento fue la colocación del dispositivo.

Este sería instalado en un vehículo que la empresa Lestido utiliza como vehículo de alquiler, por lo que la instalación debía ser lo menos visible y manipulable posible, buscando evitar accidentes de "este usuario se subió al vehículo y se enredó con el cable y ya no reporta". Como se ve en la Figura 3.17, el dispositivo queda escondido detrás y debajo del volante, fuera del campo visual del usuario, pero más importante, lejos de los pedales, evitando interferir con la conducción.



Figura 3.17: Instalación del Dispositivo SG de Smartway en Volkswagen e UP

Debido a la ubicación del mismo, le costaba captar la señal GPS para reportar correctamente los datos. Es por esto que se agrega de manera discreta una antena GPS. Con esto, el dispositivo queda funcionando por el plazo pactado con la empresa de dos meses. Aquí, en la Figura 3.18 se presenta un diagrama de la instalación final del dispositivo en el vehículo.



Figura 3.18: Diagrama del Dispositivo SG instalado en Volkswagen e UP

4. Segunda etapa: Análisis de datos

4.1. Metodología aplicada

Para el análisis de datos obtenidos del vehículo eléctrico, se implementó una metodología estructurada en cuatro fases principales: exploración y pre-procesamiento de los datos recopilados, entrenamiento del modelo de predicción de SoH y análisis de patrones de degradación. Este enfoque permite no solo identificar y limpiar los datos obtenidos a través de la telemetría del vehículo, sino también desarrollar modelos que permitan anticipar la evolución de su batería en función de su uso real. La metodología aplicada busca equilibrar la rigurosidad del análisis técnico con la necesidad de generar resultados aplicables en la gestión del estado de salud de las baterías de los vehículos.

En primer lugar, se llevó a cabo una etapa de exploración y pre-procesamiento de datos, donde se implementaron técnicas de segmentación para identificar ciclos de carga y descarga, así como la eliminación de datos anómalos y la estandarización de las series temporales. Posteriormente, se procedió con la selección y entrenamiento de un modelo de predicción de SoH basado en redes neuronales LSTM, el cual se entrenó utilizando bases de datos públicas de la NASA y el MIT y se validó utilizando los datos del vehículo.

Finalmente, el análisis de patrones de degradación permitió evaluar la evolución del estado de salud de la batería en función de sus condiciones de operación, identificando tendencias y correlaciones clave. Se exploraron diversas estrategias para la interpretación de los resultados, incluyendo el uso de valores Shapley para determinar la importancia relativa de cada parámetro en la predicción de la degradación.

4.2. Exploración de los datos recopilados

Una vez instalado el dispositivo, se registró su telemetría desde el 24 de octubre al 16 de diciembre de 2024. Los parámetros registrados en este período fueron los siguientes:

- Corriente de la batería de alto voltaje (A)
- Voltaje de la batería de alto voltaje (V)
- Temperatura de la batería de alto voltaje ($^{\circ}\text{C}$)
- Capacidad de la batería de alto voltaje (kW/h)
- Estado del vehículo
- Tipo de carga (AC o CC)
- Velocidad de conducción (km/h)
- SOC (%)
- Consumo 100km (kW)
- Kilometraje (km)

Las telemetrías fueron reportadas a un *endpoint* de la empresa Smartway donde fueron almacenadas por el periodo del proyecto. A su vez, Smartway facilitó el acceso a su plataforma web donde se puede visualizar y acceder a estos datos. Debido a que este almacenamiento no contaba con una API, los datos históricos fueron extraídos directamente desde la aplicación web de forma manual.

Como primer paso para el análisis, se analizaron los datos obtenidos mediante gráficas de las telemetrías en el tiempo.

En primer lugar, se analizó cuál fue la distribución de estados del vehículo en el tiempo, para identificar aproximadamente cuánto se utilizó el vehículo eléctrico en el período registrado.

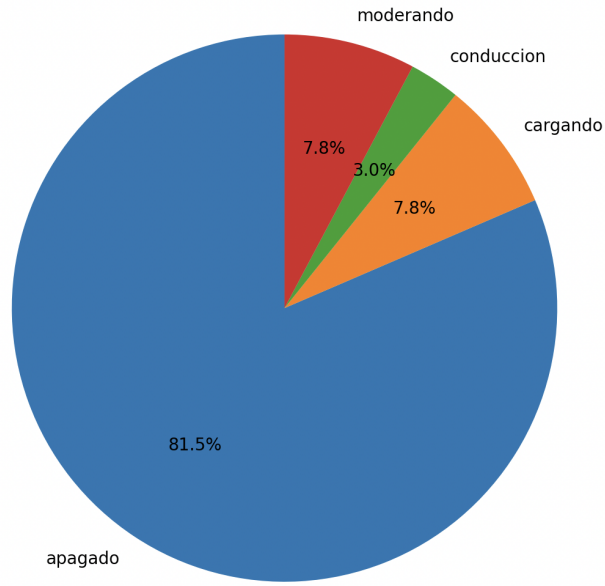


Figura 4.1: Porcentaje de utilización del vehículo

Se puede observar que, a pesar de ser un vehículo de renting, este pasó la mayor parte del tiempo apagado. Esto causa que los datos extraídos sean relativamente acotados. En este periodo, el auto recorrió un total de 987 km, que de acuerdo con la autonomía del vehículo registrada por el fabricante de 260 km, correspondería a aproximadamente 4 ciclos de carga totales bajo la definición de ciclo con profundidad de descarga total, expuesta en la sección 2.3.3

4.2.1. Pre-procesamiento de datos

En cuanto al pre-procesamiento de los datos, uno de los principales desafíos fue la identificación de ciclos de carga y descarga. Fue de vital importancia seccionar la telemetría en ciclos consecutivos de carga y descarga dado que estos presentan características distintas en el uso y cargado de las baterías. Por ejemplo, la corriente es positiva cuando el vehículo se está cargando y negativa cuando éste se está descargando. Por este motivo, se consideró que tiene valor agregado analizarlas por separado.

Esto se realizó utilizando el estado del vehículo como indicador, que recorre la telemetría ordenada cronológicamente punto por punto, utilizando los cambios de estado del vehículo para identificar ciclos. Se define que la batería se está descargando cuando el auto está

en estado moderando o conducción. Por otro lado, se define que la batería está cargando únicamente cuando el vehículo esta en estado "cargando".

Se realizó un *script* donde se implementó una máquina de estados que detecta transiciones entre estados de carga y descarga, y guarda los periodos de carga y descarga:

- La transición a estado de carga, ocurre cuando el estado del vehículo es cargando y estaba previamente descargando. Se registra el tiempo final del ciclo de descarga anterior.
- La transición a estado de descarga ocurre cuando el estado del vehículo no es cargando ni apagado. Se registra el tiempo final del ciclo de carga anterior y se guarda el ciclo en la lista de carga.

Los tiempos donde inicia y termina cada periodo fueron utilizados para seccionar las series temporales de todas las variables y así obtener una lista de ciclos seccionada con las telemetrías correspondientes a cada uno. Como resultado de este pre-procesamiento se obtuvo un total de 23 ciclos de carga y 23 ciclos de descarga.

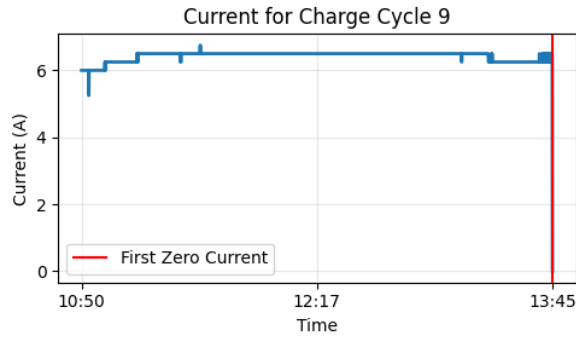
4.2.2. Depuración de los datos

En este proceso, se encontraron varios aspectos de los datos que debían ser depurados para poder ser estandarizados. Con este fin, solo se consideraron los periodos de al menos 1 minuto consecutivo para ser agregados a la lista. Por el mismo motivo, las series temporales de los periodos de carga fueron recortadas cuando la corriente de las mismas pasa a ser cero, dado que en muchas ocasiones el auto permanecía enchufado, detectando estado cargando, incluso después de haber alcanzado la carga completa.

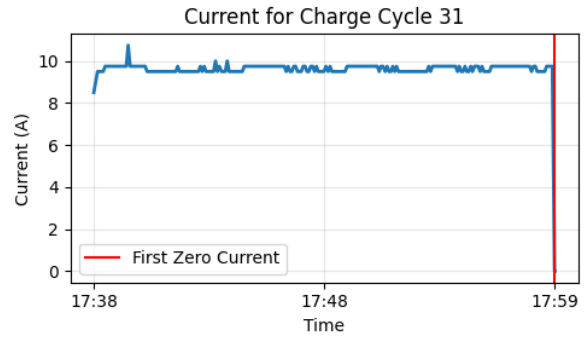
4.2.2.1 Ciclos de carga

Se analizaron las gráficas de las señales recopiladas para entender en mayor profundidad cómo son estos datos. En primer lugar, se graficaron las curvas de corriente para los ciclos de carga.

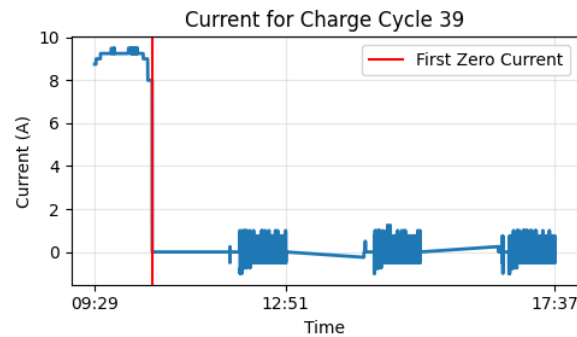
En su mayoría las curvas halladas se asemejaban a las dos primeras, donde la corriente era aproximadamente estable en torno a un valor el total del tiempo de carga. Se intuye que el vehículo se cargó utilizando dos cargadores distintos dado que en algunos ciclos de carga,



(a)



(b)



(c)

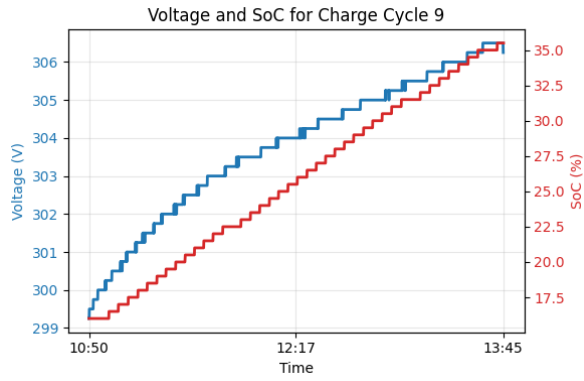
Figura 4.2: Gráficas de corriente para ciclos de carga

la corriente se estabiliza en torno al valor de 6A 4.2a mientras que en otros la corriente se estabiliza en torno al de 10 A 4.2b. Sin embargo, también se identificaron con una curva a primera vista distinta como en 4.2c.

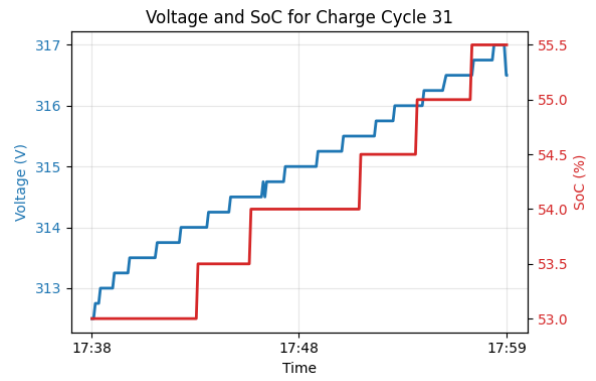
Se denotó que estos ciclos anómalos de carga tenían una extensión particularmente larga, como por ejemplo el de la figura donde dura aproximadamente de 8 horas. A su vez, se identificó que la primera parte de la gráfica coincide en forma con el resto de los diagramas realizados. El hecho de que la corriente se vuelva cero de forma repentina, da a entender que la batería del vehículo no está recibiendo más energía. Por ese motivo, se interpretó que corresponde a un momento donde la carga de la batería se completó y a pesar de eso, continuó enchufado.

En cuánto a las gráficas de voltaje se observó una tendencia similar en todos los ciclos de carga donde el voltaje de la batería aumenta a medida que transcurre la carga.

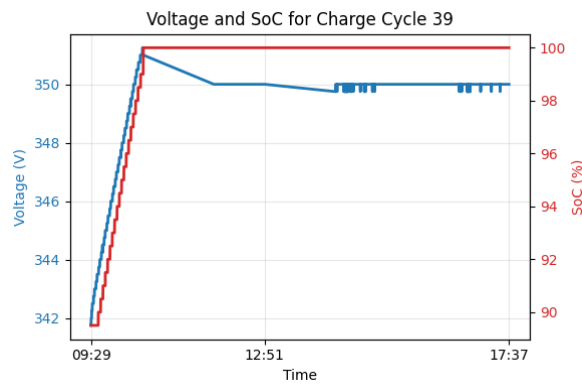
Se dedujo que este valor está relacionado con el SoC del ciclo dado que se observa que a medida que el porcentaje de carga es mayor, también lo es el voltaje. Cabe destacar, que



(a)



(b)



(c)

Figura 4.3: Gráficas de corriente para ciclos de carga

aunque se observa una estrecha correlación, esta relación es claramente no lineal. Se puede visualizar en 4.3c que el voltaje toma su valor máximo al completar la carga. Este punto es conocido como el voltaje en circuito abierto, dado que cuando se completa una carga la corriente es nula como se observó en 4.2c.

Al ser un caso borde, distinto en forma al resto, se optó por descartar la porción de la carga después que el SoC llega a 100%. Para ello, se encontró el tiempo en el cuál la corriente cruza 0A y se redefinió ese ciclo hasta ese punto. En las gráficas de corriente 4.2 este punto se denota con una línea roja perpendicular al eje del tiempo.

Por último, también se graficó la temperatura en los ciclos de carga para entender cómo varía. Todas las temperaturas registradas en ciclos de carga se encuentran entre 17°C y 29°C.

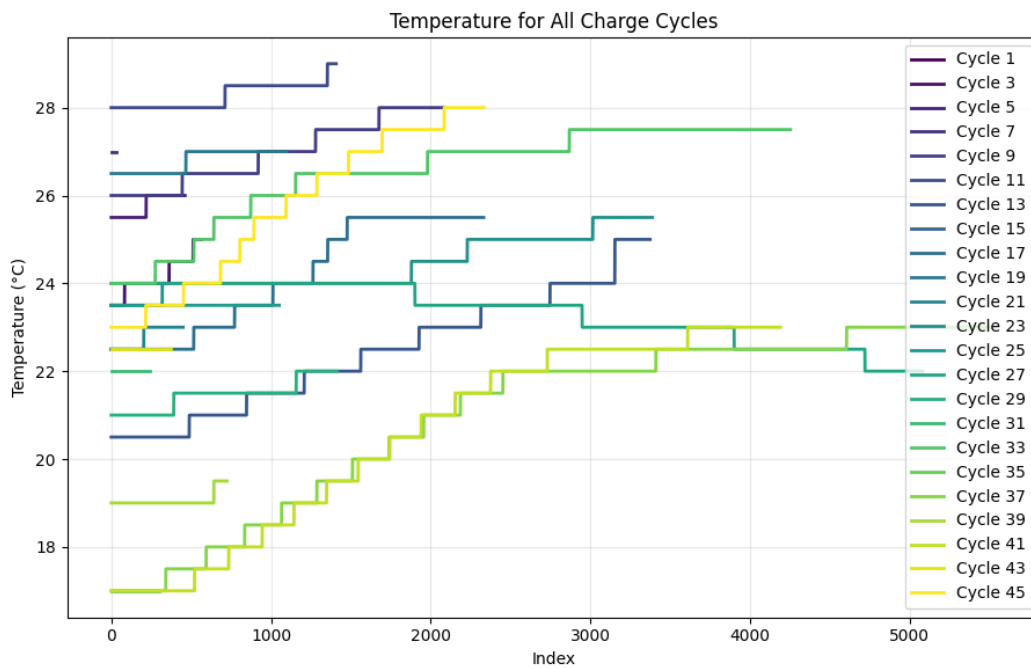


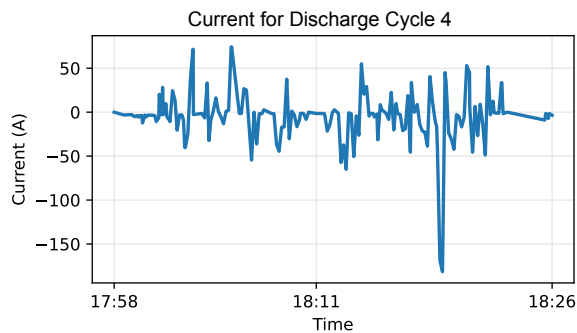
Figura 4.4: Variación de la temperatura en los ciclos de carga

Como se denota en 4.4, en la mayor parte de los ciclos se observa que la temperatura aumenta a medida que la carga transcurre.

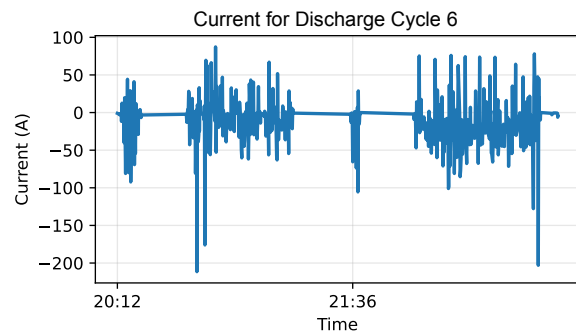
4.2.2.2 Ciclos de descarga

Para los ciclos de descarga también se graficaron las variables principales del problema para entender su dinámica en mayor profundidad.

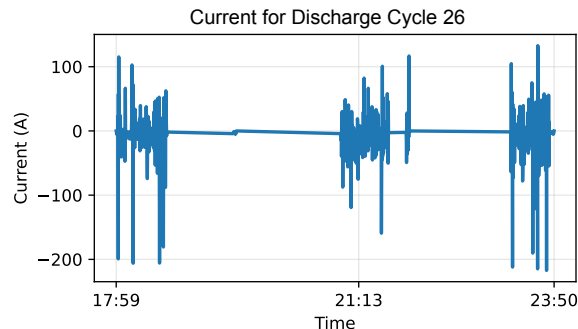
En primer lugar, se graficó la corriente de descarga para entender cómo varía en los distintos ciclos. Se puede observar que en todos los casos la corriente toma valores tanto positivos como negativos. Dado que se están analizando ciclos de descarga, los valores de corriente son más erráticos comparado a los observados en los ciclos de carga. Esta forma de onda es producto de que el auto en el tránsito no se encuentra siempre en movimiento constante. Esto genera que se presenten picos positivos de corriente cuando el vehículo frena, indicando que ingresa corriente a la batería, y picos negativos cuando el vehículo acelera, debido a que el vehículo consume corriente de la misma. Esto se debe al fenómeno de regeneración en el frenado, que fue explicado en mayor detalle en la sección 2.4.5. A su vez, se observan periodos donde la corriente se mantiene constante en un valor cercano a cero, lo que indica que el auto no está consumiendo energía y por lo tanto, está apagado o en contacto.



(a)



(b)



(c)

Figura 4.5: Gráficas de corriente para ciclos de carga

De forma análoga a los ciclos de carga, se graficó también la tensión y el SOC en los ciclos de descarga.

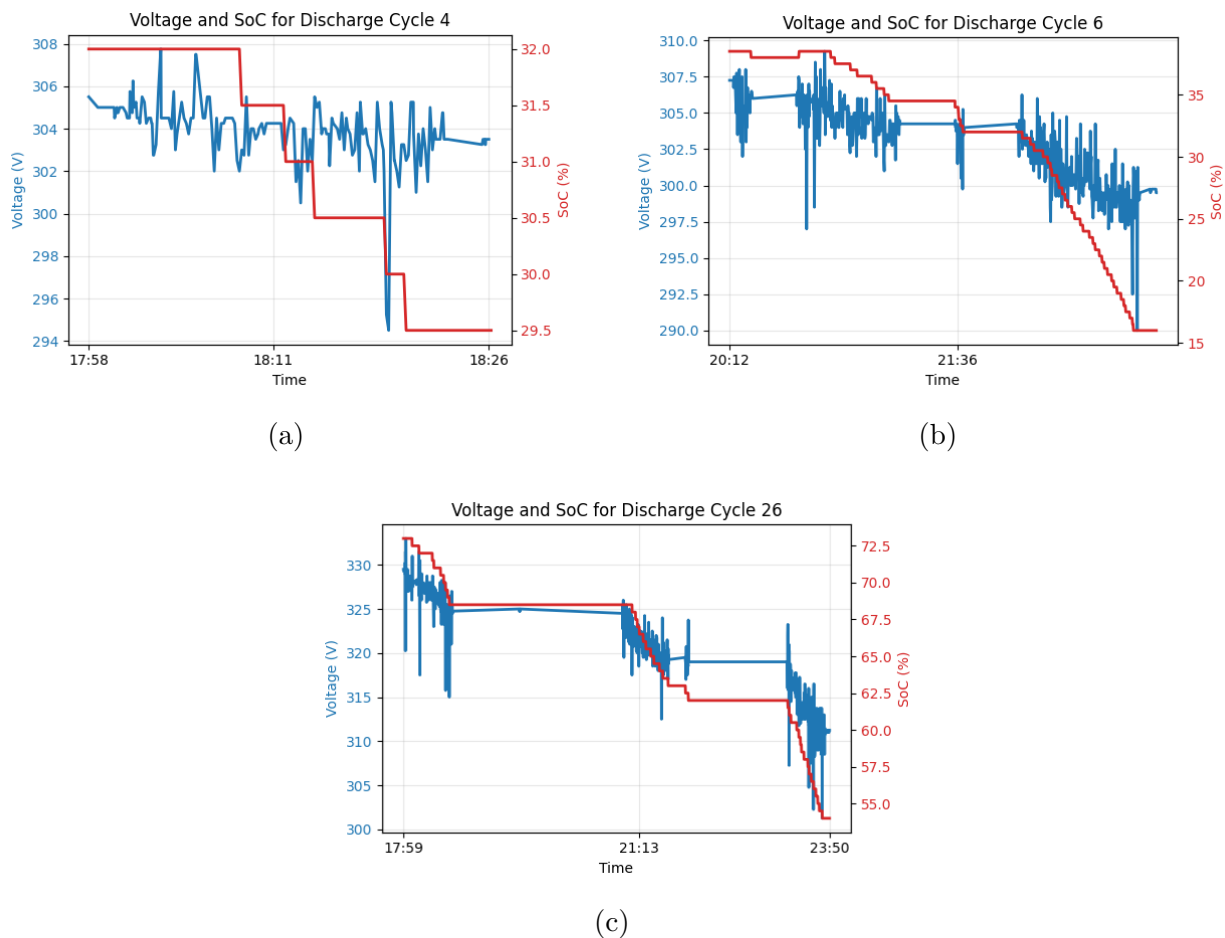


Figura 4.6: Gráficas de tensión para ciclos de descarga

Se observa que en todos los casos el SOC disminuye a medida que transcurre el tiempo, lo que es esperable dado que se están descargando.

También se observa que la tensión disminuye a medida que transcurre el tiempo, aunque no de forma monótona como en los ciclos de carga. Esto se evidencia mejor en las gráficas 4.6b y 4.6c dado que estas presentan una profundidad de descarga mayor.

De forma similar a la corriente, la tensión presenta un comportamiento fluctuante, con presencia de picos y caídas. Esto se puede atribuir a los patrones de conducción del auto en el tránsito. Estos picos y caídas coinciden en gran medida con el comportamiento de la corriente, dando a entender que la impedancia interna de la batería es mayormente resistiva.

La potencia es una variable fundamental en el análisis de baterías, ya que representa la cantidad de energía transferida por unidad de tiempo $P(t) = V(t) \cdot I(t)$. En el contexto de los vehículos eléctricos, la potencia entregada por la batería es crucial para determinar el rendimiento del vehículo, especialmente durante aceleraciones y subidas. La potencia instantánea se calcula como el producto de la corriente y la tensión, y su variación puede proporcionar información valiosa sobre el estado de salud de la batería. Para los casos graficados, se observa que la potencia toma el mismo signo de la corriente, dado que la tensión es positiva en todos los casos. Una potencia negativa indica que la batería está entregando energía al vehículo, mientras que una potencia positiva indica que la batería está siendo cargada. Las variaciones bruscas de la potencia indican cambios en la demanda de energía por aceleraciones y frenados.

Por último, se graficó la temperatura de la batería. En todos los ciclos de descarga, la temperatura se mantiene entre 17°C y 29°C, de forma similar a los ciclos de carga. A continuación se presentan los gráficos de temperatura y corriente para los ciclos de descarga 28, 40 y 42.

En estas se observa que la temperatura aumenta a medida que el auto se utiliza más, es decir en el periodo donde la corriente es distinta de cero. De la misma forma, la temperatura desciende cuando la corriente es cero, es decir cuando el auto está apagado o en contacto. En el caso del ciclo 40 4.7b se observa que la temperatura desciende a pesar de estar en conducción. Considerando que en ese caso, el auto se encuentra en una zona de alta temperatura considerando el umbral de temperaturas observado, esta observación se puede deber a que el BMS del auto activa el enfriamiento de la batería.

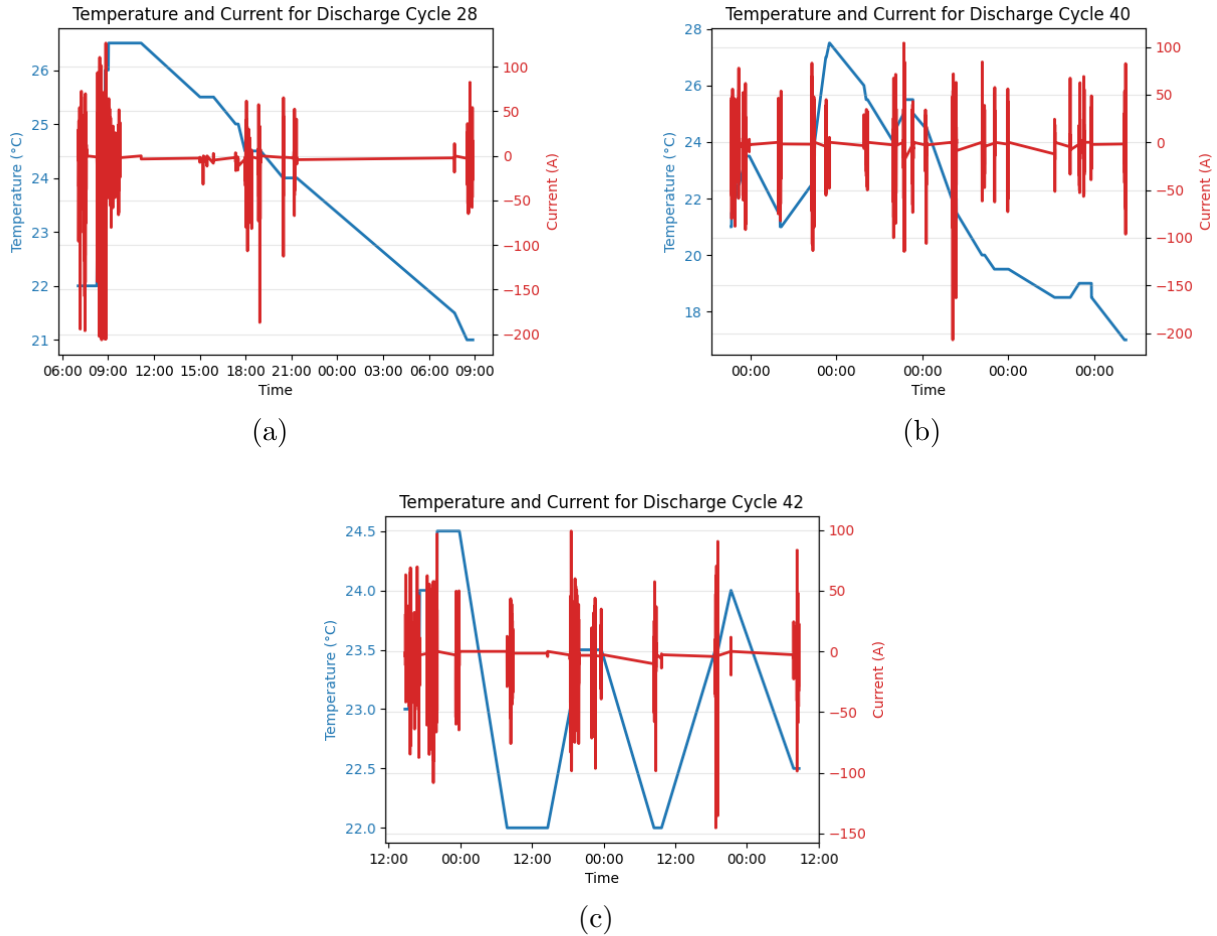


Figura 4.7: Gráficas de temperatura para ciclos de descarga

4.2.3. Estimación de la vida útil de la batería

En el marco de este trabajo, se elige el parámetro de SoH como indicador de la vida útil de la batería dado que este permite cuantificar la degradación de la batería en términos de su capacidad y por lo tanto, entender la evolución de la batería a lo largo del tiempo.

Como fue descrito anteriormente en la sección 2.3.4, el SoH se define como la proporción de la capacidad de la batería que aún puede ser utilizada en relación a su valor nominal de fábrica.

$$SoH = \frac{C_{actual}}{C_{inicial}} \times 100\%$$

En este sentido, el SoH es una variable que no puede ser medida, sino que debe ser estimada. Esta estimación puede ser realizada de distintas formas como se exploró en la sección 2.3.4.1. En algunos casos, los fabricantes de vehículos incorporan estrategias para estimar el SoH de la batería en la ECU correspondiente a el BMS del auto. Sin embargo, en el caso de este trabajo, no se cuenta con esta información directa del auto por lo que es necesario realizar una estimación del SoH a partir de los datos recopilados.

Para poder realizar el cálculo es necesario conocer la capacidad nominal de la batería, la cual se puede obtener del fabricante o midiendola al comienzo de la vida útil de la batería si se tiene la posibilidad, y la capacidad actual de la batería, siendo esta última la que se desea estimar. Para el caso de este trabajo, se exploraron dos métodos distintos para la estimación.

En primer lugar, se implementó el método de estimación de SoH a partir de las series temporales de capacidad y SOC de la batería, haciendo uso de la telemetría del vehículo. La serie temporal de capacidad corresponde al valor instantáneo de la misma en kWh, proveída por el BMS del vehículo. Se comenzó por graficar la serie temporal de capacidad en función del SOC para observar su evolución conjunta.

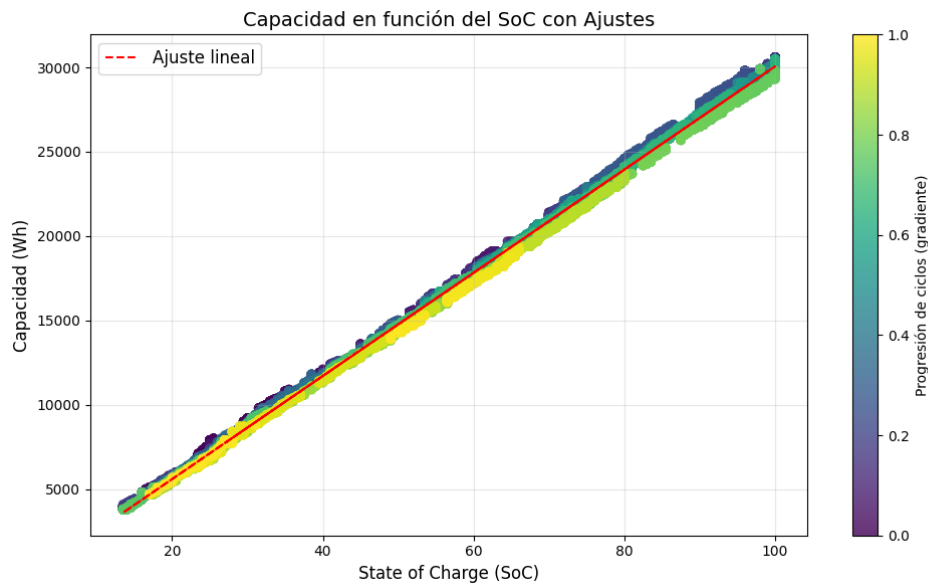


Figura 4.8: Serie temporal de capacidad en función del SOC

Cada ciclo es representado con círculos del mismo color, y se implementa un gradiente de colores para representar la progresión de los ciclos. Los primeros ciclos registrados se representan con colores más oscuros, mientras que los últimos ciclos se representan con colores más claros.

En la gráfica 4.8 se observa que la capacidad se ajusta linealmente en relación al SOC. A su vez, analizando la gráfica detalladamente, se observa que las pendientes de estas rectas va disminuyendo a medida que transcurren los ciclos.

Dado que el objetivo es estimar la capacidad total para cada ciclo, se decidió ajustar una recta a los datos de la gráfica 4.8 para cada ciclo y obtener la capacidad para un SOC del 100%. Con ese valor de capacidad, se calculó el SoH para cada ciclo como el cociente entre la capacidad actual y la capacidad nominal de la batería. De esta forma, se obtienen los valores de SoH para cada ciclo que se grafican en la figura 4.9.

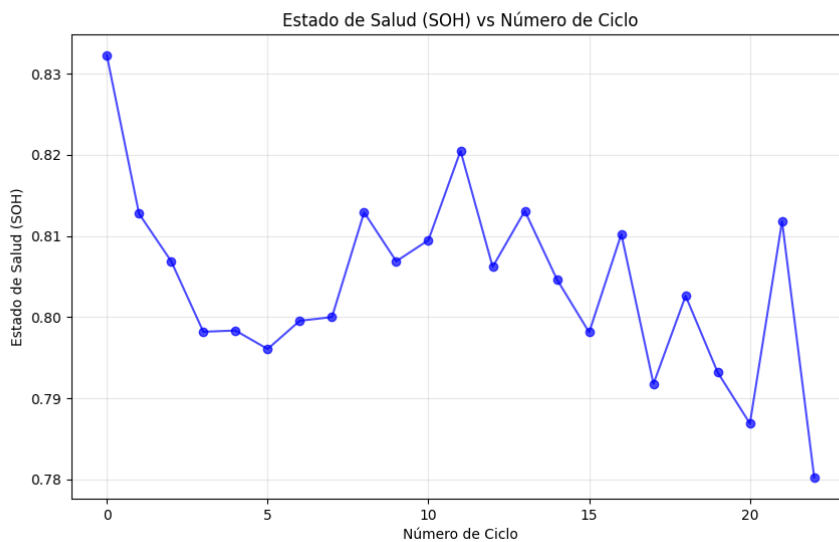


Figura 4.9: SoH vs número de ciclo

En segundo lugar, se optó por explorar el método de estimación denominado *Conteo de Coulomb*, el cual se basa en la integración de la corriente de la batería a lo largo del tiempo. En la documentación del vehículo eléctrico intervenido proveída por Lestido, el dato de la capacidad nominal de la batería se encontraba expresado en kWh en vez de amperios-hora (Ah), siendo esta la capacidad energética de la batería. Por lo tanto, se decidió adaptar el método de Coulomb estudiado en la sección 2.3.4.1 para que pueda ser aplicado a este caso. Para ello, se integró la potencia de la batería a lo largo del tiempo, obteniendo así la energía total entregada por la batería.

Una limitación de este método en cuanto a su utilización para estimar la capacidad energética de forma *on-line* es que, teóricamente, requiere que la integral se realice sobre una carga o descarga completa, comenzando desde el 0% de carga hasta el 100% de carga. Sin

embargo, como los datos obtenidos del vehículo provienen de su uso real, no todos los ciclos son completos.

Por este motivo, se decidió adaptar el método de estimación de Coulomb, integrando la potencia de la batería en ciclos parciales y ajustando el valor de la capacidad hallada en base a la profundidad de descarga de cada ciclo. Por ejemplo, en el caso del ciclo 26 de carga, 4.6 se observa que la profundidad de descarga fue de un 20 %. Para calcular la capacidad energética para ese ciclo, integramos la corriente y el voltaje y realizamos dividimos entre 0,2 para obtener la capacidad energética total correspondiente al 100 % de carga en la batería. Este método asume que la capacidad energética de la batería es directamente proporcional a la profundidad de descarga. Esto presenta una desventaja dado que la relación entre estas variables no es exactamente lineal, por lo que puede llegar a introducir ruido en el cálculo del SoH. Sin embargo, en el método anterior se evidencio de que la capacidad y el SoC estaban relacionadas casi linealmente.

$$DoD = |SoC_{final} - SoC_{inicial}|$$

$$E_{actual} = \frac{100}{DoD} \cdot \int_0^t I(t) \cdot V(t) dt$$

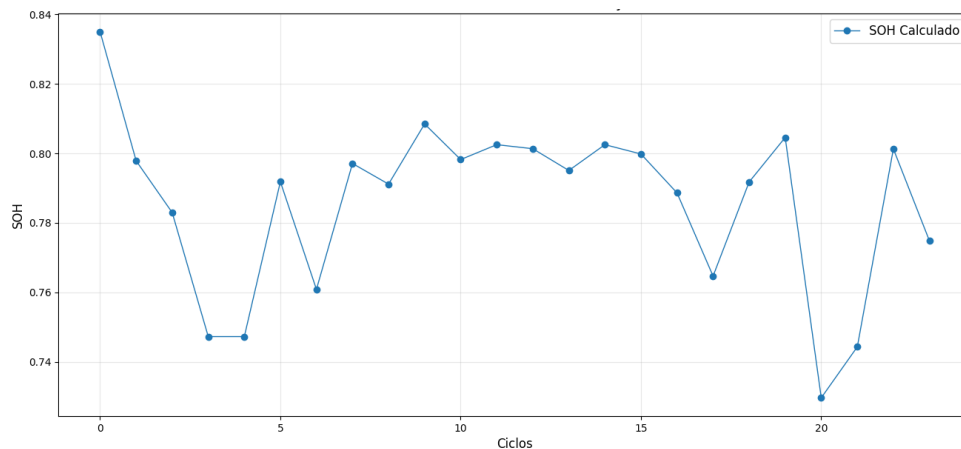


Figura 4.10: SoH vs número de ciclo

En la práctica, factores como la temperatura, la tasa de descarga y la edad de la batería pueden afectar esta relación lineal. Por ejemplo, a bajas temperaturas, la capacidad efectiva de la batería puede disminuir, lo que significa que el SoC no reflejaría con precisión la energía

disponible.[60] Para evidenciar esto, se graficó el SoH obtenido y la temperatura promedio de cada ciclo.

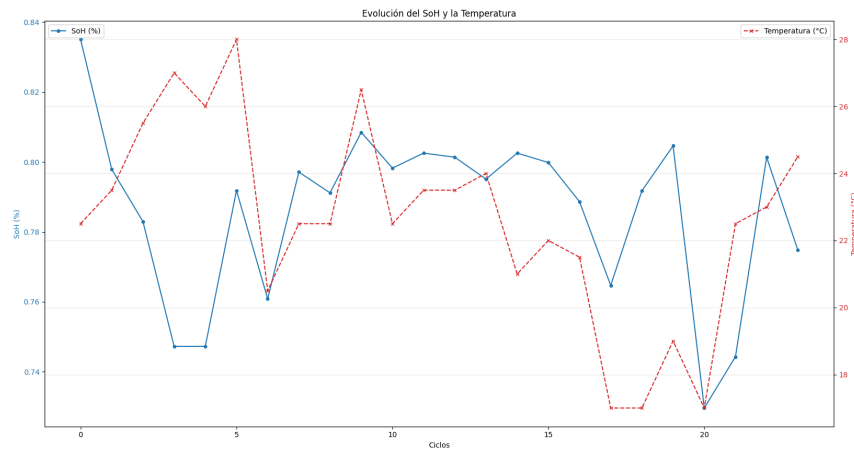


Figura 4.11: SoH vs temperatura promedio

En la gráfica 4.10 se observa que los SoHs calculados están mucho más dispersos que los obtenidos con el método anterior, tomando valores entre 0,73 y 0,84. Por este motivo, se puede deducir que este segundo método acarrea mayor error que el, probablemente debido a las incertidumbres en la medición de la corriente y el voltaje. El hecho de que la temperatura promedio en cada ciclo varíe sustancialmente como se evidencia en 4.11 da a entender que las no linealidades desatadas por variaciones en la temperatura pueden estar afectando el cálculo de la capacidad energética de la batería.

En resumen, se eligió el método de estimación de SoH a partir de la serie temporal de capacidad y SOC dado que presentó un mejor ajuste a los datos y una menor dispersión en los valores obtenidos.

4.3. Predicción de curva de degradación de SoH

4.3.1. Métodos de predicción

Estimar los parámetros dinámicos de las baterías eléctricas como el SoH es una tarea compleja por diversas razones explicadas anteriormente. En la bibliografía consultada se encontraron una serie de estrategias evaluadas para conseguir este propósito, que reiteradamente se caracterizan en dos grandes grupos: enfoques *data-driven* y enfoques *physics-based*.

Los enfoques basados en modelos físicos buscan establecer un modelo matemático que describa la trayectoria de degradación en función de la dinámica de la batería. Este enfoque se basa en la modelización física del comportamiento de degradación o en la construcción de un modelo empírico que represente la evolución del sistema a lo largo del tiempo. En general, el modelo se define mediante un conjunto de ecuaciones algebraicas y diferenciales, o bien a partir de una ecuación empírica. Una característica clave de este enfoque es que el modelo se diseña específicamente para un sistema determinado. Entre ellos, se destacan los modelos exclusivamente eléctricos como los modelos de circuito equivalente y los modelos electroquímicos que incorporan dinámicas de los procesos internos de la batería.

Por el contrario, los enfoques basados en datos emplean teorías estadísticas o técnicas de aprendizaje automático para derivar un modelo predictivo directamente a partir de datos medidos, sin necesidad de construir un modelo físico del sistema. A diferencia del enfoque basado en modelos, el enfoque basado en datos es más flexible y aplicable a distintos sistemas, ya que el mismo algoritmo puede utilizarse tanto en la predicción de la vida útil de celdas de baterías como para sistemas multi-celda como los encontrados en vehículos eléctricos.

Estos modelos han sido ampliamente utilizados en la literatura para la estimación del estado de salud y la vida útil remanente de las baterías, cada uno con diferentes niveles de precisión y requerimientos computacionales. El estado del arte en el rubro propone profundizar en los métodos *data-driven*, dado que permiten modelar relaciones más complejas y sutiles que podrían ser despreciadas en los métodos tradicionales. Con el creciente y acelerado desarrollo del área de aprendizaje automático, nuevas arquitecturas han surgido como las redes neuronales y el *Deep Learning*, que no solo permiten identificar patrones complejos sino que también generalizar el conocimiento adquirido.

Por este motivo, se decidió explorar este camino para poder realizar predicciones sobre el SoH futuro y así poder cuantificar y anticipar la degradación de las baterías de los autos eléctricos.

Los modelos basados en datos a utilizar emplean información operativa de un sistema en un momento dado para prever su comportamiento futuro. En particular, en el rubro de la predicción de SoH de baterías de iones de litio, los modelos se clasifican en tres categorías principales [61]:

1. Modelos autorregresivos no lineales (NAR)

Estos modelos predicen valores futuros de una serie temporal basándose únicamente en sus valores pasados. Son ampliamente utilizados en el monitoreo de vida útil de baterías de iones de litio debido a la tendencia de degradación suave y descendente.

2. Modelos regresivos no lineales con Variables Exógenas (NRX)

Estos modelos predicen valores futuros de una serie temporal utilizando únicamente variables exógenas, sin considerar los valores pasados de la serie en cuestión.

3. Modelos autorregresivos no lineales con variables exógenas (NARX)

A diferencia de los NAR, estos modelos incorporan variables exógenas adicionales, como voltaje, corriente y temperatura, para mejorar la precisión de las predicciones además de la serie histórica de la variable a predecir.

En la literatura consultada, múltiples estudios detallan que existe una relación entre las variables exógenas como la corriente, voltaje y temperatura sobre el SoH. A su vez, es de nuestro interés entender como los distintos patrones de uso y del entorno del vehículo afectan la vida útil de su batería. Por este motivo, se decidió ir por un modelo del tipo NARX dado que nos pareció el que mejor se ajustaba a los objetivos del proyecto.

Algunas implementaciones de este tipo de modelo utilizan la resistencia interna de las baterías, sin embargo, este parámetro es difícil de medir de forma *on-line* y como nuestro objetivo es que el modelo pueda ser utilizado sobre cualquier vehículo sin tener la necesidad de agregar sensores invasivos, este no va a ser tenido en cuenta.

En resumen, la aplicación de modelos basados en datos, especialmente aquellos que emplean redes neuronales, ha demostrado ser efectiva en la predicción del comportamiento y la degradación de baterías de iones de litio. Se eligió un modelo del tipo NARX donde

las entradas al modelo son datos de temperatura, voltaje, corriente y SoH para cada ciclo, considerando los datos que se pudieron extraer exitosamente del vehículo eléctrico intervenido.

4.3.2. Conjuntos de datos utilizados

Los métodos de predicción de SoH basados en datos dependen del acceso a un conjunto grande de datos extraídos y validados previamente de los cuales se puedan entrenar los modelos. Para el entrenamiento de nuestros modelos, se utilizaron dos conjuntos de datos públicos.

4.3.2.1 Conjunto de datos de NASA

El primer conjunto de datos utilizado, publicado por el Centro de Excelencia en Prognósticos de la NASA Ames, se recopiló utilizando un banco de pruebas personalizado para pronósticos de baterías. Fue extraído de [62]. En este estudio, las baterías de iones de litio fueron sometidas a tres perfiles operativos diferentes: carga, descarga y espectroscopía de impedancia electroquímica, bajo diversas temperaturas. Las descargas se realizaron con diferentes niveles de carga de corriente hasta que el voltaje de la batería alcanzó umbrales preestablecidos, asociados al voltaje de corte inferior de cada batería. Los ciclos repetidos de carga y descarga resultaron en un envejecimiento acelerado de las baterías, y los experimentos se detuvieron cuando las baterías alcanzaron el criterio de fin de vida útil, definido como una disminución del 30 % en la capacidad nominal.

Los datos del mismo vienen dados en un archivo donde se distinguen ciclos de tres categorías: carga, descarga e impedancia. Se decidió solo poner foco en los ciclos de carga y descarga, dado que en los ciclos de impedancia se realiza un ensayo de espectroscopía de impedancia electroquímica (EIS). Esto se justifica dado que el objetivo del proyecto es llegar a un modelo con el cual se pueda hacer predicciones de SoH de forma *online*, por lo cual no sería posible tener estos datos de forma no invasiva.

Cada uno de estos ciclos contiene una tabla con las series temporales de tiempo, voltaje, corriente y temperatura medidos en los terminales de la batería y sus análogos, pero medidos en los extremos del cargador o la carga para los ciclos de carga y descarga, respectivamente. Como se puede apreciar en la figura 4.12, la carga se realiza en dos etapas, una de corriente constante y otra a continuación de voltaje constante. Por otra parte, en la figura 4.13 se muestra que la descarga se realiza a una corriente constante de 2 A. Se resalta que se sigue

la misma notación de signo que los datos observados en el vehículo real en la sección Pre-procesamiento de datos, siendo la corriente negativa cuando la batería es descargada y positiva cuando es cargada.

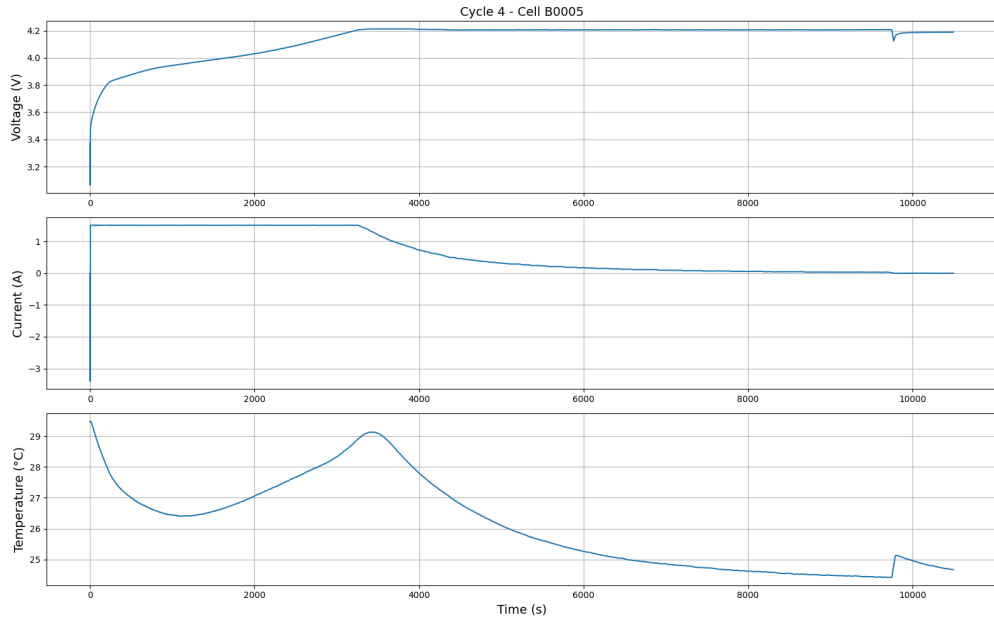


Figura 4.12: B005 - Ciclo de carga

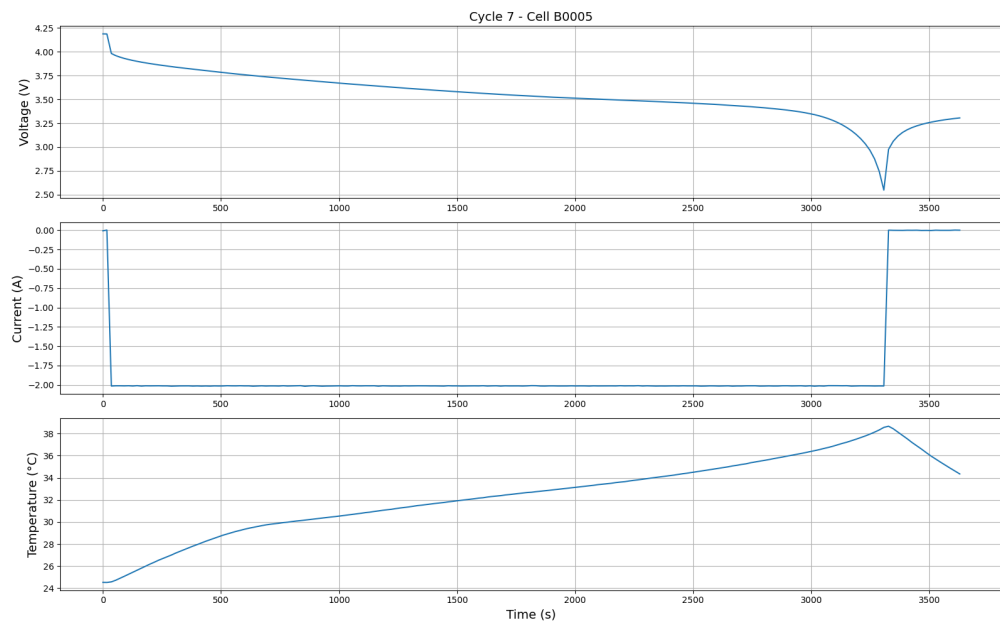


Figura 4.13: B005 - Ciclo de descarga

A su vez, exclusivamente para los ciclos de descarga se cuenta con un valor asociado a la capacidad de la batería en ese ciclo. El SoH por ciclo se deduce a partir de este valor, normalizando la capacidad entre su valor máximo, que en todos los casos se obtiene en el ciclo inicial.

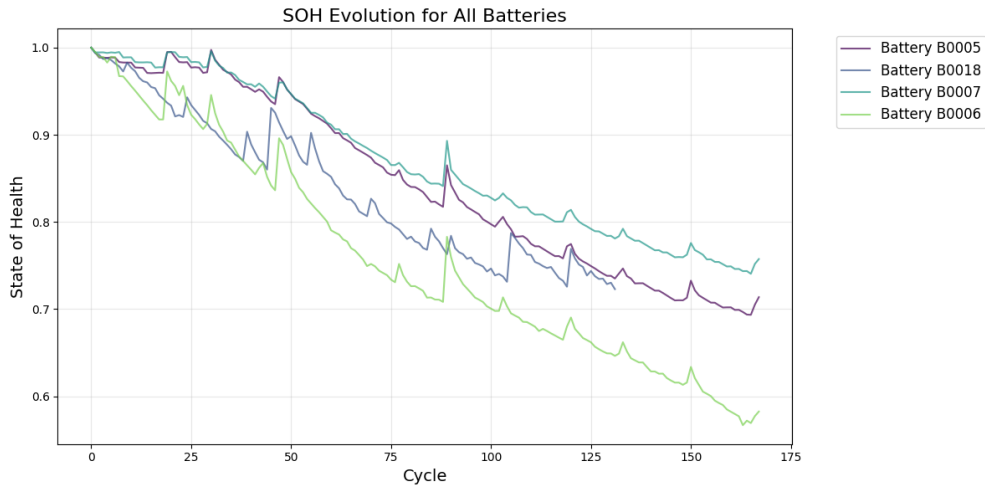


Figura 4.14: Evolución de SoH para cada celda

4.3.2.2 Conjunto de datos de MIT

El segundo conjunto de datos fue publicado en 2019 por el MIT en colaboración con *Toyota Engineering* y el Departamento de Ciencia e Ingeniería de Materiales de la *Universidad de Stanford*, introducido en el artículo [63]. Este conjunto de datos es el más grande disponible en relación con el envejecimiento de baterías de iones de litio. Las celdas utilizadas para las pruebas son celdas de fosfato de hierro y litio (LFP) con ánodo de grafito del fabricante A123, modelo APR18650M1A. Estas celdas tienen un voltaje nominal de 3.3 V y una capacidad nominal de 1.1 Ah. Las pruebas se realizaron en una cámara climática a 30°C, donde las baterías se descargaron constantemente a una corriente de 4.4 A (4C). Un factor clave en las pruebas fue la política de carga: las baterías se cargaron siguiendo un protocolo de corriente constante/voltaje constante (CC–CV) en múltiples etapas.

En él se realizaron pruebas sobre 140 celdas, en las cuales eran cargadas completamente y luego descargadas. Se procedió a visualizar las curvas de las señales analizadas, siendo estas voltaje, corriente y temperatura. Se eligió la gráfica 4.15 para representar al conjunto de datos en esta explicación; sin embargo, se pueden encontrar otras gráficas pertinentes en el anexo Visualización de los datos MIT.

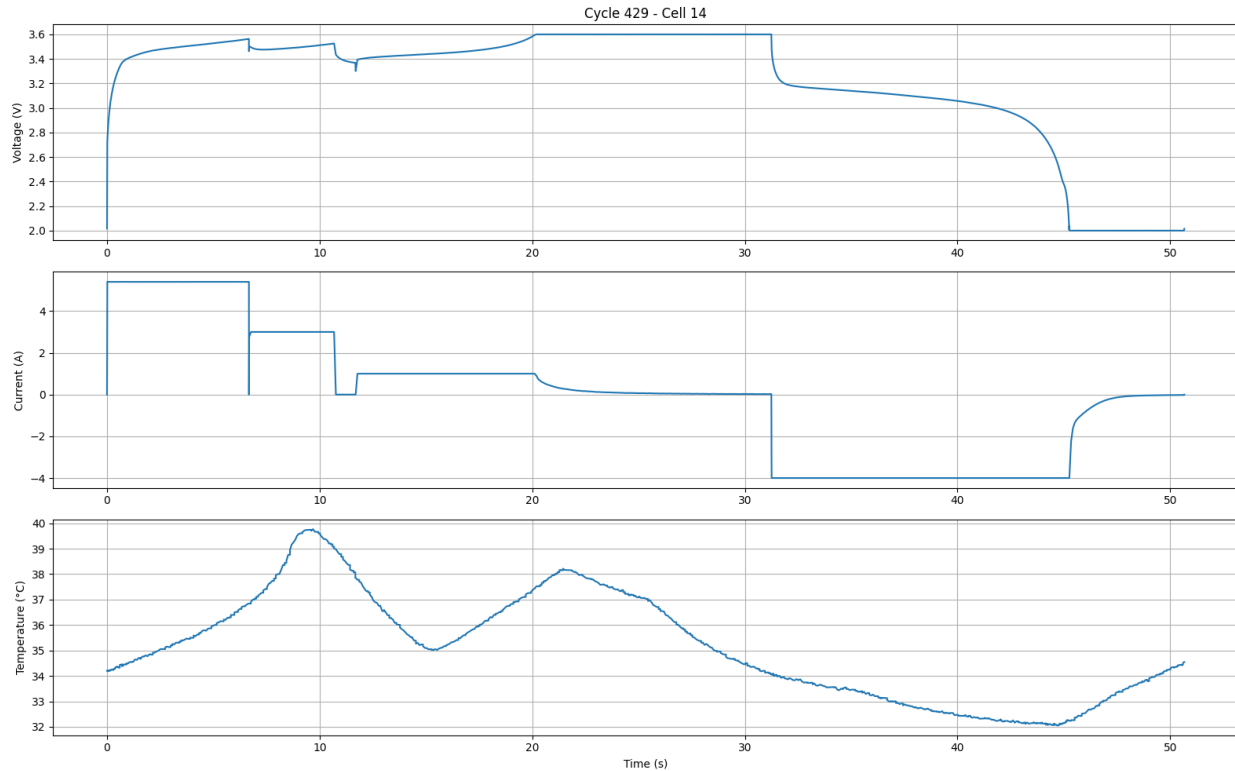


Figura 4.15: Gráficas de voltaje, corriente y temperatura para la Celda 14 - Ciclo 429

En primer lugar, se denota que, a diferencia del conjunto de datos de NASA, cada serie temporal contaba con los datos concatenados de la sección de carga y descarga. Esto se puede visualizar en la figura 4.15, identificando el cambio de signo que se manifiesta en la gráfica de corriente. Por este motivo, hubo que realizar un pre-procesamiento de los datos para separar ambas secciones y así obtenerlas en el mismo formato que los datos de NASA donde los ciclos de carga y descarga se consideraban por separado.

Analizando las curvas de carga para múltiples ciclos, se observa los distintos protocolos de carga utilizados, con distintas magnitudes y duraciones de las secciones constantes.

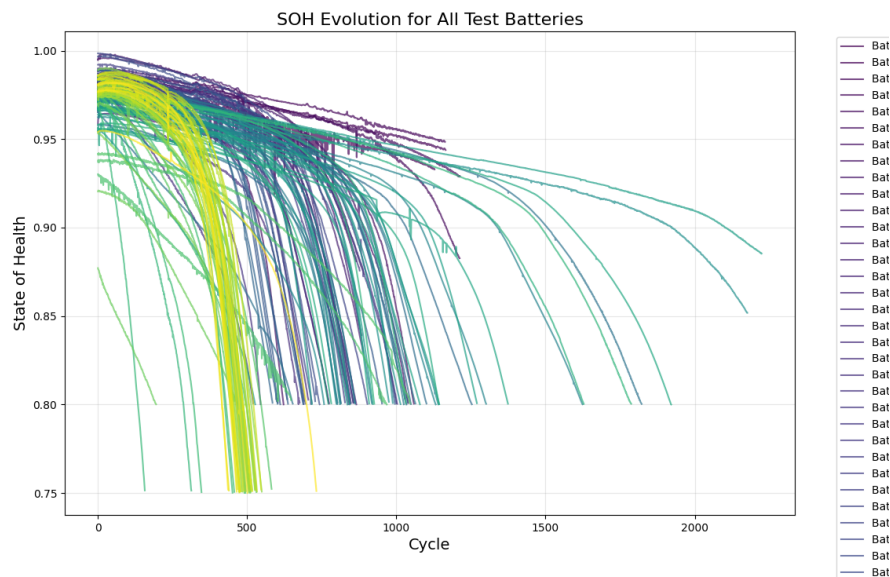


Figura 4.16: Evolución del SoH por celda - MIT

En la figura 4.16, se encuentran graficadas las curvas de SoH para todas las celdas del conjunto de datos. Se observa que la mayoría las curvas graficadas tienen un punto de inflexión, a partir del cual la capacidad comienza a disminuir de manera más pronunciada. A su vez, se denota que en parte de los experimentos se consideró el fin de vida útil llegar a un SoH de 80 % mientras que para otras muestras fue de 75 %. Es de importancia evaluar que la mayor cantidad de ciclos registrados se corresponden a SoH mayores a 90 %, alineados con el hecho de que este parámetro disminuye de forma menos pronunciada al comienzo de la vida útil de la batería.

4.3.2.3 Comparación de los conjuntos de datos

la combinación de los conjuntos de datos de NASA y MIT ha permitido entrenar y validar modelos de estimación de SoH bajo condiciones de operación muy diversas, aunque ambos difieren en diversos aspectos técnicos que condicionan su utilización.

En primer lugar, la química de las celdas y las condiciones de prueba son claramente distintas. Mientras que NASA ensayó baterías de iones de litio con perfiles de carga/descarga y mediciones de impedancia electroquímica a distintas temperaturas, el conjunto de MIT se basó en celdas LFP. Considerar distintas químicas de baterías es favorable para este estudio dado que en el caso de los vehículos eléctricos, en la mayoría de los casos se desconoce la

composición química de sus baterías dado que son propietarias. Por este motivo, se considera valioso poder analizar como se comportan baterías con diferencias en su química interna.

En segundo lugar, se observa un rango de ciclos significativamente distinto en cada base de datos. Los ensayos de NASA, como se ve en la figura 4.20, contemplan típicamente hasta unos 125–175 ciclos, con un perfil de degradación progresivo que refleja el envejecimiento acelerado. Por el contrario, en MIT (figura 4.21) se registran algunas celdas con más de 2000 ciclos, evidenciando una primera fase de degradación lenta seguida de un descenso más marcado tras un punto de inflexión. Este contraste se acentúa por los protocolos de carga/descarga utilizados: NASA implementa distintos niveles de corriente y umbrales de voltaje para proteger las celdas, mientras que MIT recurre a un esquema de carga CC–CV en múltiples etapas y una descarga constante a 4.4 A (4C).

4.3.3. Selección de arquitectura del modelo

La selección de la arquitectura del modelo para la predicción del estado de salud (SoH) de baterías de ion-litio se fundamentó en un análisis de los enfoques más utilizados en la literatura y en la comparación de su desempeño en problemas similares. Para ello, se consideraron diversas metodologías de predicción de series temporales, dado que el modelo elegido debe poder interpretar sucesiones de datos de entrada prestando especial atención al orden de los mismos.

Se analizaron una serie de arquitecturas basadas en aprendizaje profundo, como redes neuronales recurrentes (RNN), unidades recurrentes cerradas (GRU), redes de memoria a largo plazo (LSTM), redes convolucionales (CNN) y modelos basados con transformadores. Cada una de estas arquitecturas presenta ventajas y limitaciones según el tipo de datos y la complejidad del problema. Si bien los modelos basados en transformadores han demostrado un desempeño prometedor en la predicción de series temporales, su alto costo computacional y la necesidad de grandes volúmenes de datos dificultan su implementación práctica en este contexto. Por otro lado, las redes CNN han sido utilizadas para extraer características relevantes en datos secuenciales, pero su capacidad para capturar dependencias de largo plazo es limitada. En contraste, los modelos recurrentes como RNN, GRU y LSTM han mostrado una mayor capacidad para procesar información secuencial con dependencias temporales a largo plazo, lo que los convierte en candidatos más adecuados para este tipo de predicción.

Además del análisis teórico de los modelos, se llevó a cabo una revisión de la literatura con el objetivo de identificar enfoques previos aplicados a la predicción del SoH en baterías. En este sentido, estudios como el de Khumprom y Yodo [64] han demostrado la efectividad de los modelos basados en redes neuronales profundas en la estimación del SoH y la vida útil restante (RUL) de baterías de ion-litio. En particular, estos trabajos destacan que los modelos de aprendizaje profundo superan a enfoques tradicionales como máquinas de soporte vectorial (SVM) y regresión lineal en términos de precisión y generalización. Por otro lado, investigaciones más recientes, como la de Lv et al. (2024), han explorado arquitecturas híbridas que combinan redes convolucionales y LSTM para mejorar la representación de características espaciales y temporales en los datos de baterías. Sin embargo, estos modelos presentan una mayor complejidad computacional, lo que puede dificultar su implementación en entornos con recursos limitados. A partir de esta revisión, se observó que la arquitectura LSTM se menciona reiteradamente como una de las alternativas más eficientes para este problema, debido a su capacidad de capturar patrones temporales sin incurrir en la complejidad de modelos más avanzados.

La decisión de utilizar un modelo basado en LSTM se basó en tres factores fundamentales. En primer lugar, estas redes permiten capturar dependencias de largo plazo en los datos, lo que resulta esencial para la predicción del SoH, ya que la degradación de las baterías ocurre de manera acumulativa a lo largo del tiempo. A diferencia de las RNN convencionales, las LSTM incorporan mecanismos de memoria que mitigan el problema del desvanecimiento del gradiente, lo que les permite aprender de manera más efectiva la evolución del estado de la batería. En segundo lugar, se buscó un equilibrio entre complejidad computacional y desempeño predictivo. Aunque los modelos más sofisticados, como los transformadores o las combinaciones de CNN y LSTM, han demostrado mejoras en la precisión, su implementación requiere un mayor costo en términos de procesamiento y almacenamiento de datos. En cambio, las LSTM han mostrado ser una alternativa más eficiente, con un rendimiento competitivo sin necesidad de incrementar excesivamente la carga computacional. Finalmente, en pruebas comparativas realizadas con diferentes arquitecturas [65], el modelo LSTM presentó los mejores resultados en términos de error de predicción en comparación con RNN y GRU, y obtuvo un desempeño similar al de enfoques más complejos, lo que reafirma su idoneidad para esta tarea.

En conclusión, la arquitectura LSTM fue elegida por su capacidad para modelar eficazmente la evolución del estado de salud de las baterías a lo largo del tiempo, proporcionando una solución robusta y computacionalmente eficiente para la predicción del SoH.

4.3.4. Introducción a LSTM

Las redes de memoria a largo corto plazo (Long Short-Term Memory, LSTM) son una arquitectura específica de las redes neuronales recurrentes (RNN) diseñada para abordar las limitaciones de las RNN tradicionales en el manejo de dependencias a largo plazo y el problema del desvanecimiento del gradiente. Introducidas por Hochreiter y Schmidhuber en 1997 [66], las LSTM incorporan una estructura de memoria interna que permite retener información relevante durante períodos prolongados, facilitando el aprendizaje en tareas secuenciales complejas.

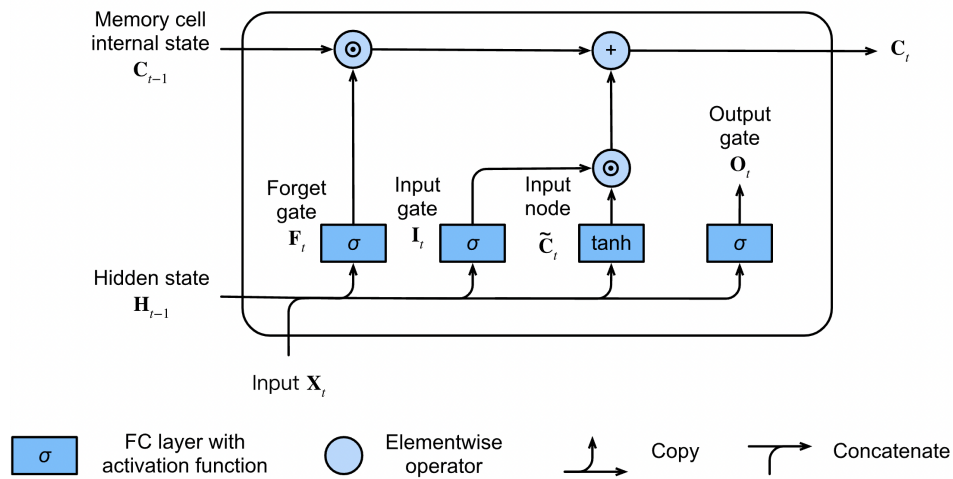


Figura 4.17: Cálculo del estado interno de una celda en un modelo [18]

4.3.4.1 Descripción de la celda LSTM

Una celda LSTM consta de varios componentes que interactúan para gestionar el flujo de información:

- Estado de la celda (Memory cell internal state): Actúa como una vía de transporte que recorre toda la secuencia, permitiendo que la información fluya con mínimas modificaciones de una celda a otra. En el diagrama se denota como C_t .
- Puerta de entrada (Input Gate): Controla qué información nueva se almacenará en el estado de la celda, evaluando la importancia de la información entrante. En el diagrama se denota como I_t .

- Puerta de olvido (Forget Gate): Decide qué información del estado de la celda debe ser descartada, evaluando qué datos no son relevantes para la predicción actual. En el diagrama se denota como F_t .
- Puerta de salida (Output Gate): Determina qué parte del estado de la celda se utilizará para generar la salida en el paso temporal actual. En el diagrama se denota como O_t .

Estas puertas trabajan en conjunto para regular el flujo de información dentro de la celda LSTM, permitiendo que la red aprenda y retenga dependencias temporales de manera efectiva.

4.3.4.2 Formato de entrada al modelo LSTM

En lo que respecta al formato de entrada a un modelo LSTM, es fundamental comprender varios aspectos clave que determinan su efectividad en el procesamiento de datos secuenciales:

Los modelos LSTM requieren que los datos de entrada mantengan una estructura secuencial explícita, donde cada fila o instancia representa un paso temporal específico en una secuencia ordenada. Este orden cronológico permite que el modelo capture dependencias temporales y patrones evolutivos en los datos. A diferencia de otros algoritmos de aprendizaje automático, las LSTM son sensibles al orden temporal, lo que significa que la permutación de las instancias alteraría completamente la interpretación y el aprendizaje del modelo.

Debido a la arquitectura secuencial de los modelos LSTM, no requieren que las series temporales sean de la misma longitud. Sin embargo, en el artículo [61] se propone un modelo que emplea ventanas de entrada pequeñas que abarcan varios ciclos consecutivos en lugar de toda la información desde el inicio de la vida de la batería. Este enfoque ofrece mayor flexibilidad al modelo, elimina la necesidad de almacenar grandes secuencias de datos y simplifica el preprocesamiento al trabajar con ventanas de tamaño fijo. De esta manera, el modelo puede ser entrenado offline y realizar predicciones en tiempo real sobre datos nuevos no vistos previamente por el modelo en cualquier punto de la vida de la batería.

Para transformar las series temporales obtenidas en un formato más conciso, frecuentemente la técnica de ventana deslizante. Esta metodología consiste en segmentar la serie temporal original en múltiples subconjuntos superpuestos, cada uno compuesto por una secuencia de observaciones consecutivas. Cada ventana se desplaza progresivamente a lo largo de la serie, avanzando típicamente un paso temporal en cada iteración. Formalmente, si denotamos

la longitud de la ventana como n , en cada paso se consideran los valores $x_t, x_{t+1}, \dots, x_{t+n-1}$ para predecir uno o más valores futuros. En el marco de este proyecto, a cada ventana se le asocia un valor de SoH futuro. La elección del tamaño de ventana constituye un hiperparámetro crítico que debe equilibrar la captación de patrones a largo plazo con la eficiencia computacional y el riesgo de sobreajuste. Esta técnica permite transformar el problema de predicción de series temporales en un formato supervisado convencional, donde cada ventana actúa como una instancia de entrenamiento.

En el contexto de las series temporales utilizando ventanas deslizantes, se calculan *features* o características que representan las variables que se utilizan como entrada al modelo LSTM. Estos pueden derivarse directamente de las mediciones originales o construirse mediante técnicas de ingeniería de características [67]. Entre los métodos más comunes para la extracción de características en series temporales se encuentran:

- Características estadísticas: Incluyen medidas como la media, desviación estándar, percentiles, asimetría o curtosis calculadas sobre ventanas temporales.
- Características de dominio de frecuencia: Obtenidas mediante transformaciones como la Transformada Rápida de Fourier (FFT), que capturan componentes periódicos y oscilaciones en diferentes escalas temporales.
- Características en el dominio del tiempo: Variables que representan información temporal como la derivada, la integral y la energía.

4.3.4.3 Descripción de la arquitectura elegida

La arquitectura implementada para la predicción del SoH se basa en una red neuronal recurrente LSTM (Long Short-Term Memory) de dos capas, siguiendo el enfoque propuesto por Jorge et al. [61] en su trabajo sobre extracción de características de series temporales para la predicción del estado de salud de baterías de ion-litio. La primera capa LSTM contiene 64 unidades con retorno de secuencias habilitado, seguida por una capa de Dropout del 10% para prevenir el sobreajuste. La segunda capa LSTM, más densa, incorpora 256 unidades, conectándose finalmente a una capa de salida densa con una única neurona para la predicción del valor de SoH. El modelo se optimiza mediante el algoritmo Adam con una tasa de aprendizaje de 0.001.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 5, 64)	19,456
dropout (Dropout)	(None, 5, 64)	0
lstm_1 (LSTM)	(None, 256)	328,704
dense (Dense)	(None, 1)	257
Total params: 348,417 (1.33 MB)		
Trainable params: 348,417 (1.33 MB)		
Non-trainable params: 0 (0.00 B)		

Figura 4.18: Arquitectura del modelo LSTM implementado

Dado que la cantidad de puntos en los *datasets* utilizados principalmente contenían datos para SoH mayores a 0,9, se decidió utilizar una función de pérdida de error cuadrático medio ponderado (weighted MSE), que asigna mayor importancia a las predicciones menos frecuentes, es decir aquellas cuyo SoH objetivo es menor a 0,9, mejorando así la precisión en la estimación de la degradación futura de la batería. Esta configuración permite capturar eficientemente tanto las dependencias temporales a corto plazo como los patrones de degradación a largo plazo presentes en las series de datos de las baterías.

4.3.5. Pre-procesamiento de los *datasets*

Previo al entrenamiento del modelo, es necesario preprocesar los datos para que sean entradas adecuadas para el mismo. Como fue detallado en la sección anterior, esto implica:

1. Cálculo y selección de las características que serán utilizadas como variables de entrada del modelo.
2. Aplicación de técnicas de normalización sobre los datos de entrada.
3. Definición de hiperparámetros como el tamaño de la ventana deslizante y el horizonte temporal objetivo para la predicción del modelo.

4.3.5.1 Definición de features

Como fue profundizado anteriormente 4.3.2, los *datasets* encontrados cuentan con información de múltiples celdas, para cada una de ellas la información se secciona en ciclos de carga y descarga y cada uno de estos ciclos contiene series temporales de voltaje, corriente y temperatura.

Cabe destacar, que no todas las celdas tienen la misma cantidad de ciclos de carga y descarga por lo que la longitud de las series temporales es variable. A su vez, las series temporales de voltaje, corriente y temperatura dentro de cada ciclo tienen diferente cantidad de puntos. Por otro lado, este modelo se va a utilizar para hacer inferencia sobre ciclos de carga de autos reales, por lo que es necesario que el modelo pueda generalizar sobre ciclos de diferente longitud.

Según lo analizado en 4.3.4, las dimensiones del tamaño del modelo LSTM son fijas. Esto implica que la cantidad de features para cada observación y el tamaño de la ventana deben ser constantes.

Por este motivo, se decidió aplicar ingeniería de características, mejor conocido como *feature engineering* para reducir las series temporales de voltaje, corriente y temperatura a una cantidad finita de valores por ciclo. Analizando el estado del arte en la literatura [61], se encontró que las características usualmente calculadas para este tipo de aplicaciones:

- **Dominio estadístico:**

- *Media*: el valor promedio de la señal.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

- *Raíz Cuadrática Media (RMS)*: la raíz cuadrada del promedio de los cuadrados de los valores

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

- *Máximo*: el valor máximo de la señal.

$$x_{max} = \text{máx}(x_1, x_2, \dots, x_n)$$

- *Mínimo*: el valor mínimo de la señal.

$$x_{min} = \text{mín}(x_1, x_2, \dots, x_n)$$

- **Dominio de frecuencia:** Transformada Rápida de Fourier (FFT), tomando la frecuencia fundamental no nula, frecuencia del máximo armónico y el ancho de banda de potencia.

- *Frecuencia Fundamental* ($f_{fundamental}$): la frecuencia correspondiente al valor máximo en el espectro de la FFT:

$$f_{fundamental} = f_k \quad \text{donde} \quad k = \arg \text{máx}_j |X_j|$$

Donde:

- X_j es el valor de la FFT en la frecuencia f_j
- f_k es la frecuencia donde la amplitud es máxima
- *Frecuencia Máxima* (f_{max}): la frecuencia máxima en el espectro.

$$f_{max} = \text{máx}(f_1, f_2, \dots, f_{n/2})$$

- *Ancho de Banda de Potencia (PBW)*: la frecuencia que contiene el 95% de la potencia de la señal.

$$PBW = f_m \quad \text{donde} \quad m = \text{mín}\left\{j : \sum_{i=1}^j |X_i|^2 \geq 0,95 \cdot \sum_{i=1}^{n/2} |X_i|^2\right\}$$

Donde:

- X_i es el valor de la FFT en la frecuencia f_i
- f_m es la frecuencia en el índice m

- **Dominio temporal:**

- *Energía Total* (E_{total}): la energía total de la señal normalizada por el tiempo total.

$$E_{total} = \frac{\sum_{i=1}^n x_i^2}{t_n - t_1}$$

Donde:

- x_i es el valor de la señal en el punto i

- t_n es el tiempo final
- t_1 es el tiempo inicial
- n es el número total de muestras
- *Área Bajo la Curva AUC*: calculada mediante la formula de área del trapecio.

$$AUC = \sum_{i=1}^{n-1} \frac{(t_{i+1} - t_i) \cdot (x_i + x_{i+1})}{2}$$

Donde:

- x_i es el valor de la señal en el tiempo t_i
- x_{i+1} es el valor de la señal en el tiempo t_{i+1}

4.3.5.2 Selección de *features*

En total, tenemos tres señales que se presentan en los tres conjuntos de datos, siendo estas el voltaje, la corriente y la temperatura. Para cada una de estas, se realizó el cálculo de las *features* en los tres dominios anteriormente mencionados de forma de poder representar las principales características de las curvas temporales en un vector unidimensional. Sin embargo, no todas las propiedades calculadas aportan el mismo valor para todas las señales. Por lo que se procedió a hacer una selección de estas previo al entrenamiento. Se probaron distintos conjuntos de características con el fin de obtener la combinación de las mismas que minimizara el error RMS.

En primer lugar, se hizo referencia a un artículo dentro de la literatura consultada donde se realizaron pruebas con distintas combinaciones de estos y se derivó que las características más significativas fueron las siguientes [61].

- **Corriente:**
 - Energía total en carga
 - Energía total en descarga
- **Voltaje:**
 - Área bajo la curva en carga
 - Área bajo la curva en descarga
 - Ancho de banda de potencia en descarga

- Media de voltaje en descarga
- **Temperatura:**
 - Máximo en carga
 - Máximo en descarga
 - Frecuencia máxima en carga
 - Ancho de banda de potencia en descarga

Para poder utilizar como punto de comparación e inferir conclusiones propias de este estudio, se comparó el enfoque anteriormente mencionado con el más generalista que implica incluir todas las características de todas las señales en el entrenamiento. Esto se realiza bajo la premisa de que la arquitectura seleccionada debería poder identificar aquellas características con poca influencia y disminuir su participación ajustando los pesos de la matriz que las multiplica.

Además de las features calculadas a partir de las series temporales mencionadas, para cada ciclo, se incluyó como entrada el SoH actual dado que se optó por un modelo del tipo NARX.

Los resultados de los modelos entrenados serán comparados y analizados en la sección Modelos entrenados con el *dataset* del MIT.

4.3.5.3 Normalización de features

La normalización es un paso fundamental para el entrenamiento efectivo de modelos de deep learning, especialmente relevante para arquitecturas LSTM. Los modelos LSTM son particularmente sensibles a la escala de los datos de entrada debido a sus funciones de activación (sigmoide y tangente hiperbólica) y a su arquitectura recurrente que propaga gradientes a través del tiempo. La normalización contribuye a una convergencia más rápida durante el entrenamiento, reduce la varianza en las actualizaciones de los pesos y permite que el modelo asigne importancia relativa adecuada a cada característica, independientemente de su escala original.

Para ello, sobre cada característica calculada se aplicó la siguiente transformación:

$$x_{normalized} = \frac{x}{max(|x|)}$$

Donde $\max(|x|)$ representa el valor máximo absoluto de la característica x en todo el conjunto de datos.

Este enfoque de normalización asegura que todas las características se encuentren en el rango $[-1,1]$, preservando la información direccional de las variables mientras se elimina la influencia de las diferentes escalas. Los valores máximos utilizados para la normalización se almacenan en un diccionario que se guarda para su posterior uso en la fase de inferencia, asegurando que los nuevos datos sean normalizados con los mismos factores de escala, para características como la temperatura que son extrínsecas al tipo de batería.

4.3.5.4 Ventana de entrada al modelo

La entrada al modelo LSTM corresponde a una serie de pasos temporales de dimensión n , siendo n la cantidad de *features* de entrada. Hasta ahora, se tenía dos tipos de conjuntos de datos con características distintas, los datos de carga y los datos de descarga. Dado que ambos no comparten exactamente las mismas *features*, fue necesario combinar las secciones de carga y descarga en un único arreglo unidimensional para representar cada paso temporal de manera consistente.

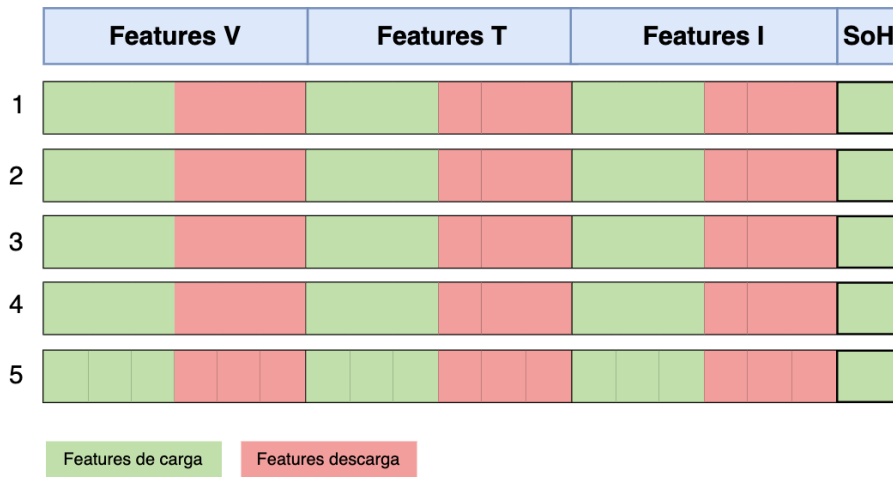


Figura 4.19: Esquemático de la ventana de entrada al modelo LSTM

Cada ventana de entrada contiene 5 pasos temporales, que se alimentan secuencialmente al modelo LSTM. En cada paso, la entrada incluye las *features* correspondientes a la sección de carga, las *features* de la sección de descarga y el SoH asociado a ese instante temporal.

4.3.5.5 Tamaño de la ventana deslizante y horizonte temporal objetivo

Parte de los objetivos de este proyecto fue predecir el SoH de la batería de un vehículo eléctrico en el futuro. Dado que el SoH es un parámetro que varía lentamente con el tiempo, este horizonte temporal debe ser lo suficientemente largo como para que pueda aportar información relevante para el usuario. A continuación, se hará referencia a este parámetro como T .

Por otro lado, la selección del largo de la ventana deslizante, W , tiene un efecto directo en la cantidad de datos históricos que el modelo puede utilizar para realizar sus predicciones. Usualmente, se asocia un tamaño mayor de la ventana con un mejor rendimiento del modelo, sin embargo, es necesario explorar con varias alternativas para encontrar el tamaño óptimo. Como detalle adicional para el criterio de selección, se tuvo en cuenta que cuanto mayor sea el tamaño de la ventana, más datos se deberán recopilar de un vehículo en particular para poder realizar una predicción.

La elección de estos parámetros estuvo condicionada por la cantidad limitada de datos disponibles para validar el modelo, ya que únicamente se cuenta con 23 ciclos de carga y descarga extraídos del vehículo. Por lo tanto, para realizar la validación del modelo utilizando n ventanas, se debe cumplir la siguiente condición:

$$W + T \leq 23 - (n - 1)$$

Distintas combinaciones de W y T fueron evaluadas en la etapa de entrenamiento para encontrar el tamaño óptimo de la ventana deslizante y el horizonte temporal objetivo. Para poder realizar una evaluación justa del modelo sobre los datos recabados, se tuvo en cuenta que el número mínimo de ventanas de validación, n , mantenga una cantidad de datos suficiente para realizar una evaluación confiable del modelo.

4.3.6. Entrenamiento del modelo

En el correr del proyecto, se realizaron varias pruebas con diferentes configuraciones de hiperparámetros para entrenar el modelo LSTM. Esto permitió encontrar el tamaño óptimo de la ventana deslizante y el horizonte temporal objetivo, así como también la mejor configuración de hiperparámetros para el modelo. También exploramos distintas estrategias en

el entrenamiento, como distintos tamaños de *batch* y de *epochs* para reducir el tiempo de entrenamiento sin afectar la precisión del modelo. A continuación, se detallan las pruebas realizadas y los resultados obtenidos.

4.3.6.1 Implementación

Para la implementación del modelo se utilizó TensorFlow como *framework* principal de desarrollo. Esta elección se fundamentó en la familiaridad del equipo con esta biblioteca y en las ventajas que ofrece para la implementación de redes neuronales recurrentes. A diferencia de otros *frameworks* populares como PyTorch, TensorFlow proporciona una API de alto nivel (Keras) que incluye capas LSTM pre-implementadas, lo que simplificó considerablemente el proceso de desarrollo. Esta característica permitió que el entrenamiento e inferencia del modelo fuera prácticamente "*plug and play*", reduciendo el tiempo de implementación y facilitando la experimentación con diferentes configuraciones de hiperparámetros.

4.3.6.2 CUDA

Para optimizar el rendimiento del entrenamiento, se utilizó CUDA (*Compute Unified Device Architecture*), una plataforma de computación paralela desarrollada por NVIDIA. CUDA permite aprovechar la capacidad de procesamiento masivo de las GPUs (Unidades de Procesamiento Gráfico), lo que resulta en una aceleración significativa de los cálculos intensivos requeridos en el entrenamiento de redes neuronales.

Las GPUs son especialmente eficientes para el procesamiento en paralelo, una característica clave en la computación de alto rendimiento y el aprendizaje profundo. Gracias a su arquitectura altamente paralelizable, las *CUDA-enabled* GPUs pueden ejecutar miles de hilos de procesamiento en paralelo, lo que reduce drásticamente los tiempos de entrenamiento en comparación con las CPU tradicionales.

Además, CUDA es compatible con varias bibliotecas de aprendizaje profundo ampliamente utilizadas, en particular con TensorFlow, lo que facilita la implementación y optimización de modelos de redes neuronales recurrentes (RNNs). Esto ha permitido al equipo acelerar el proceso de entrenamiento del modelo, manejar grandes volúmenes de datos y realizar inferencias en menor tiempo.

4.3.6.3 cuDNN

La librería *CUDA Deep Neural Network* de Nvidia es una librería de primitivas para redes neuronales profundas acelerada por GPU. Esto significa que sus funciones están optimizadas para ser ejecutadas por la GPU. Entonces, cuDNN está provista de implementaciones altamente optimizadas de operaciones comúnmente utilizadas en aplicaciones de redes neuronales profundas.[68]

Es así que TensorFlow utiliza la API de CUDA para enviar operaciones a la GPU en lugar de a la CPU y utiliza cuDNN para optimizar ciertos cálculos que son comunes en DNNs. En conjunto, TensorFlow, CUDA y cuDNN aceleran y optimizan el proceso de entrenamiento de las redes neuronales profundas.

4.3.6.4 Compatibilidad y versiones del software utilizado

Para poder utilizar CUDA y cuDNN con TensorFlow es necesario ser consciente de las distintas versiones de las mismas que se utilizan, ya que no todas las versiones de Python, TensorFlow, CUDA y cuDNN son compatibles entre sí.

Para empezar TensorFlow ya no soporta uso de GPU en computadoras Windows. Debido a que la última versión de Tensorflow que permite utilizar GPU en windows es TensorFlow 2.10, debemos basar toda nuestra instalación en función de esto.

En primer lugar, la última versión de Python compatible con TensorFlow 2.10 es Python 3.10 así que se empieza por instalarlo [69]. Luego, las versiones de CUDA y cuDNN compatibles con TensorFlow 2.10 son CUDA 11.2 [70] y cuDNN 8.1.0 [71], ambas se pueden descargar desde la página oficial de Nvidia. Una vez instalado todo lo anterior se procede a instalar Tensorflow 2.10 utilizando el comando adecuado.

Teniendo ya instalado Python 3.10, TensorFlow 2.10, CUDA 11.2 y cuDNN 8.1.0; resta solo entrenar un modelo, Tensorflow entonces detectará la GPU siempre que haya una disponible y la utilizará para el entrenamiento.

4.3.6.5 Optimización de rendimiento mediante *batch size*

En deep learning, el *Batch size* es un hiperparámetro que determina la cantidad de ejemplos de entrenamiento que se utilizan en una misma iteración. Al utilizar un valor mayor, se procesará una mayor cantidad de información en paralelo, ayudando a maximizar las

ventajas provistas por el uso de la GPU. Entrenando un modelo utilizando una GPU acelerará el proceso incluso si el valor de *Batch size* se mantiene igual, pero aumentándolo la velocidad de entrenamiento aumentará aun más gracias a la especial capacidad de paralelización de las GPUs.

Esto sin embargo tiene un límite, ya que una vez que se llega al punto en que la memoria de la GPU se llena por completo, ya no se obtendrá un mayor beneficio del incremento del *Batch size* y hasta incluso puede llegar a ser ligeramente perjudicial. Adicionalmente, valores excesivamente altos del mismo, pueden hacer que el modelo sea propenso a sobreajustar y perder precisión. En general, es clave encontrar un balance entre un valor que permita entrenar rápidamente, sin perjudicar significativamente al modelo.

Finalmente, utilizando GPU para entrenar y habiendo elegido un *Batch size* adecuado que ayude a acelerar aún mas el proceso de entrenamiento, se logró reducir el tiempo de entrenamiento de un modelo que originalmente, sin utilizar GPU, había demorado aproximadamente unas 36 horas a tan solo unos 50 minutos aproximadamente, mas de cuarenta veces mas rápido.

4.3.7. Pruebas realizadas

Se entrenaron modelos con los dos *datasets* de entrenamiento disponibles, utilizando distintos tamaños de ventana deslizante y horizonte temporal objetivo, así como también diferentes configuraciones de hiperparámetros para el modelo. Con cada uno de los modelos entrenados, se evaluó el rendimiento utilizando el *dataset* de validación, el cual fue seleccionado aleatoriamente. Se decidió no entrenar modelos con ambos datasets, ya que las características de los datos de ambos *datasets* son muy distintas, y la proporción de puntos de entrenamiento del dataset de NASA es mucho menor que el *dataset* del MIT.

4.3.7.1 Modelos entrenados con el *dataset* de NASA

Para el entrenamiento, se utilizaron 6 de las baterías proveídas por el *dataset* de NASA, siendo estas las baterías B0005, B0006, B0007, B0018, B0054 y B0055, ya que la bibliografía consultada [61] menciona que estas son las baterías que presentan datos de mayor calidad. Se implementó una estrategia de validación cruzada leave-one-out, donde en cada iteración se utilizó una batería como conjunto de prueba y las restantes como conjunto de entrenamiento. Este enfoque permitió maximizar el uso de los datos disponibles y obtener una evaluación más

robusta del rendimiento del modelo. Se estableció un número máximo de 10.000 épocas para el entrenamiento, con un mecanismo de detención temprana que monitoreaba la pérdida en el conjunto de validación con una paciencia de 100 épocas, lo que evitó el sobreajuste y optimizó el tiempo de entrenamiento. Para cada modelo entrenado, se calculó el error cuadrático medio (RMSE) sobre el conjunto de prueba como métrica de evaluación, y los modelos resultantes fueron almacenados para su posterior análisis y comparación. Esta metodología permitió identificar la configuración óptima de hiperparámetros, destacando la importancia del tamaño de la ventana deslizante y el horizonte temporal de predicción en el rendimiento final del modelo. Aplicando la estrategia de validación mencionada, se obtuvieron 6 modelos entrenados, uno para cada batería.

Se realizaron pruebas con diferentes configuraciones de tamaño de ventana deslizante y horizonte temporal objetivo para determinar la combinación óptima. Específicamente, se evaluaron ventanas de tamaño 3, 5 y 8 ciclos, con horizontes temporales de 14, 10 y 5 ciclos respectivamente.

Los resultados mostraron diferencias significativas en el rendimiento, que son comparadas a continuación evaluando el error mínimo entre los modelos entrenados con las configuraciones mencionadas. Con una ventana de 3 ciclos se obtuvo un error mínimo de 12,5 %, mientras que la ventana de 8 ciclos produjo el mejor resultado con un error mínimo de 5,5 %. La configuración con ventana de 5 ciclos mostró un rendimiento intermedio con un error mínimo de 6,6 %.

<i>Window Size</i>	<i>Temporal Horizon</i>	<i>Best RMSE</i>
3	14	12.51 %
5	10	6.68 %
8	5	5.52 %

Cuadro 4.1: Tabla comparativa de modelos realizados con *dataset* de NASA.

Como era esperado, una mayor ventana de visualización histórica captura mejor las tendencias de SoH, sin embargo, limita la cantidad de ciclos a futuro que se pueden predecir. Dado que entre las ventanas de 5 y 8 el error disminuye sutilmente, se opta por poder predecir un horizonte temporal mayor así aportando más valor a la misma. Considerando el perfil de

conducción del vehículo en cuestión, dado que se obtuvieron 23 ciclos en 2 meses, un horizonte temporal de 10 ciclos implicaría aproximadamente predecir el SoH un mes en el futuro.

En conclusión, se eligió la configuración de una ventana deslizante de 5 ciclos y un horizonte temporal de 10 ciclos, buscando minimizar el error de las predicciones pero al mismo tiempo compensando para así tener suficientes puntos de validación sobre los datos extraídos del vehículo intervenido.

4.3.7.2 Modelos entrenados con el *dataset* del MIT

Para el entrenamiento con el *dataset* del MIT, se implementó una estrategia de validación cruzada *k-fold*, dividiendo el conjunto de datos en múltiples particiones para maximizar la utilización de las muestras disponibles. Los primeros entrenamientos se realizaron utilizando el conjunto selecto de *features* descrito en la sección 4.3.5.2 dado que el tiempo en que demora en converger el algoritmo es significativamente menor que utilizando todas las características calculadas como entrada.

Este enfoque fue especialmente valioso dado el mayor volumen de datos en comparación con el *dataset* de NASA, permitiendo una evaluación más exhaustiva del modelo. A su vez, entrenar múltiples modelos permite un análisis independiente a las variaciones de resultados producto de la selección de conjunto de entrenamiento y validación.

Se mantuvieron los hiperparámetros base que resultaron óptimos en el entrenamiento con el *dataset* de NASA: una ventana deslizante de 5 ciclos, un horizonte temporal de predicción de 10 ciclos, un tamaño de lote de 32 muestras y un máximo de 10.000 épocas con detención temprana. Los datos fueron normalizados utilizando el mismo procedimiento aplicado al *dataset* de NASA, asegurando la consistencia en el preprocesamiento. Para cada *fold* de la validación cruzada, se entrenó un modelo independiente y se evaluó su rendimiento mediante el RMSE en el conjunto de prueba correspondiente.

Se comenzó entrenando utilizando la pérdida cuadrática media (MSE) como función de *loss*, pero se observó que el modelo demoraba en converger y los resultados no eran satisfactorios. Por este motivo se decidió explorar con otras funciones de pérdida. Particularmente, las predicciones inferidas del modelo entrenado con RMSE se ajustaban bien a la sección del SoH previo al punto de inflexión, pero realizaban ajustes extraños para la última parte donde el SoH decrecía con mayor pendiente. Recordando las observaciones realizadas en la exploración del conjunto de datos, se denotó que la mayor parte de los ciclos de entrenamiento tenían valores de SoH mayores a 90 %. Esto resulta coherente dado que ya se analizó que la

batería se degrada a un ritmo menor al comienzo de su vida útil. Sin embargo, esto estaba generando que el modelo ignore de cierta forma los ciclos con SoH inferiores a 90 % debido a su baja representación en el entrenamiento. Los resultados obtenidos a partir de los cuales se realizaron estas observaciones se pueden encontrar en el anexo Para mitigar este efecto, se experimentó utilizar una función de pérdida de error cuadrático medio ponderado, que asigna mayor importancia a las predicciones menos frecuentes, es decir aquellas cuyo SoH objetivo es menor a 90 %.

Se definió la función de pérdida ponderada (Weighted MSE Loss) como:

$$\mathcal{L}(y_{\text{true}}, y_{\text{pred}}) = \frac{1}{N} \sum_{i=1}^N w_i \cdot (y_{\text{true},i} - y_{\text{pred},i})^2$$

donde:

$$w_i = \begin{cases} 1, & \text{si } y_{\text{true},i} > 0,9 \\ 5, & \text{si } y_{\text{true},i} \leq 0,9 \end{cases}$$

El factor de 5 se dedujo de la diferencia de representación de ambos conjuntos a partir de la visualización de las gráficas de SoH en función del ciclo. Esta estrategia generó el efecto esperado, mejorando así la precisión en la estimación de la degradación futura de la batería.

Por último, se evaluó el desempeño del modelo al entrenar con todas las features calculadas como entrada, sin aplicar ningún tipo de filtro. Como era de esperarse, estos entrenamientos tomaron casi cuatro veces más en converger en comparación con aquellos realizados con un conjunto reducido de features. Sin embargo, los errores obtenidos al validar sobre el dataset de MIT fueron significativamente mayores que los obtenidos con el grupo selecto de features. Al analizar las gráficas, se observa que estos modelos no logran captar ni siquiera la tendencia del SoH, generando resultados aparentemente aleatorios. Esto sugiere que el modelo estaba subajustando, fenómeno conocido como underfitting. Se entiende que esto ocurrió debido a que un mayor número de características de entrada obliga al modelo a interpretar patrones en un espacio de mayor dimensionalidad, lo que dificulta la identificación de relaciones significativas en los datos. Por este motivo, se consideró como modelo final aquel entrenado con el conjunto selecto de features sugerido por el paper [61].

Como resultado del entrenamiento, se obtuvo 10 modelos entrenados, uno para cada *fold* de la validación cruzada. Cada modelo fue evaluado mediante el RMSE en el conjunto de prueba correspondiente. Se observó que los modelos tuvieron RMSE entre 0,075 % y 0,316 % y que en promedio el RMSE fue de 0,14 %. Se pueden encontrar gráficas de los resultados obtenidos en el anexo A.6.

4.3.8. Evaluación del modelo sobre datos reales

Como último paso de validación, se evaluó el modelo sobre los datos reales obtenidos del vehículo intervenido. Cabe destacar que los datos de carga y descarga de la batería en el vehículo intervenido fueron normalizados utilizando el mismo procedimiento aplicado a los datos de los *datasets* de entrenamiento, para abstraer al modelo de las claras diferencias en magnitud de tensión y corriente entre los datos de laboratorio y los del vehículo.

Para la validación, fueron utilizados los modelos obtenidos con las mejores configuraciones de hiperparámetros para cada conjunto de datos. En primer lugar, para los modelos entrenados con el *dataset* de NASA, se utilizó el entrenamiento con 5 ciclos de ventana deslizante y 10 ciclos de horizonte temporal objetivo. Se probó evaluar con los 6 modelos entrenados por separado y con los 6 modelos a modo de *ensemble*, promediando los resultados de cada modelo usando el promedio y el promedio ponderado por el inverso del RMSE. Los mejores resultados fueron obtenidos con el promedio ponderado por el inverso del RMSE. Todas las gráficas de los resultados obtenidos se pueden encontrar en el anexo A.7.

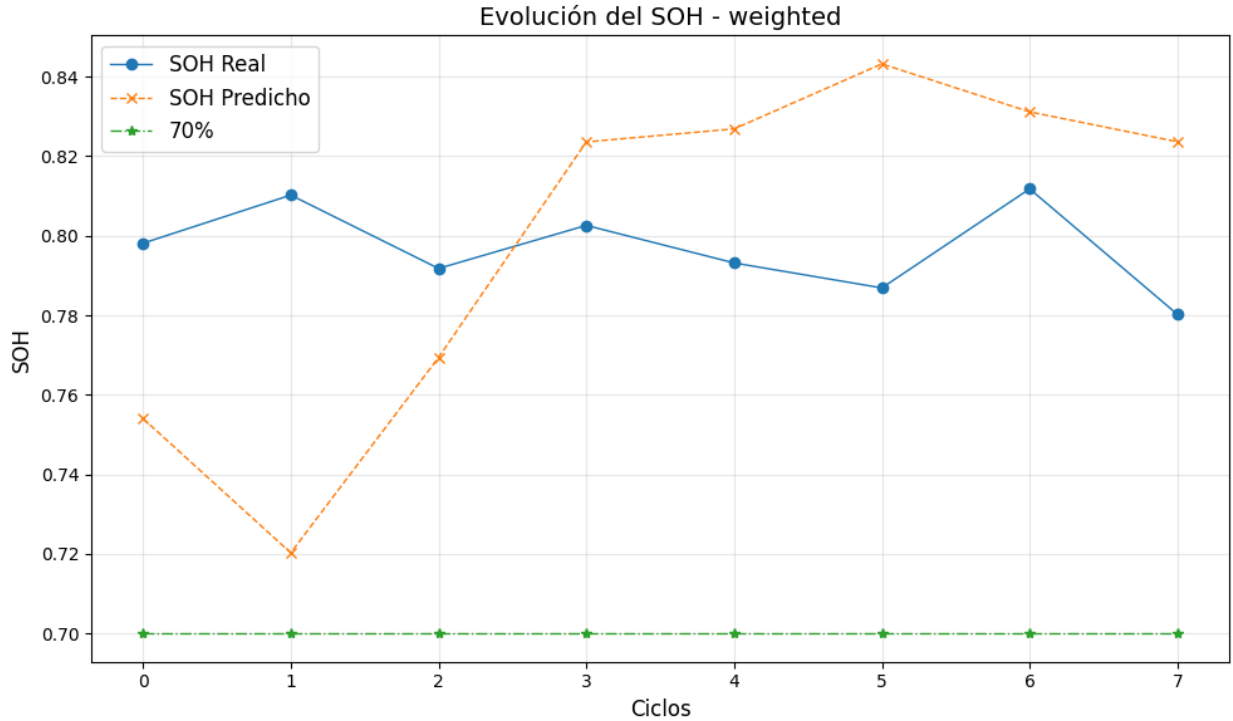


Figura 4.20: Predicciones del SoH de la batería en el vehículo intervenido utilizando el modelo NASA

Evaluando la figura 4.20 se denota que los resultados obtenidos por este modelo no son satisfactorios dado que presentan errores de casi 10% en el peor caso.

Por otra parte, para los modelos entrenados con el *dataset* del MIT, se utilizó el modelo entrenado con 5 ciclos de ventana deslizante y 10 ciclos de horizonte temporal objetivo.

El entrenamiento que tuvo mejores resultados en la validación cruzada *k-fold* con el *dataset* del MIT, fueron los obtenidos entrenando con la función de pérdida de error cuadrático medio ponderado. Se realizó la inferencia sobre los datos del vehículo utilizando los 10 modelos por separado y obteniendo valores de RMSE entre 1,0% y 2,7%. Dado que es deseable que el modelo pueda generalizar más allá de solamente el caso siendo evaluado, en vez de elegir el modelo con menor error, se definió que el modelo final fuese un *ensamble* donde el resultado de una predicción fuese el promedio de todas las predicciones individuales. Utilizando este enfoque, se logró un valor RMSE de 1,36%, siendo muy superior a los obtenidos en los modelos entrenados con el *dataset* de NASA.

También se exploró realizar un *ensamble* con predicciones ponderadas de acuerdo al error RMSE obtenido de cada modelo con su set de validación en MIT, sin embargo esto ob-

tuvo resultados peores de 2,0 %, por lo que se optó por el *ensamble* con pesos equitativamente distribuidos.

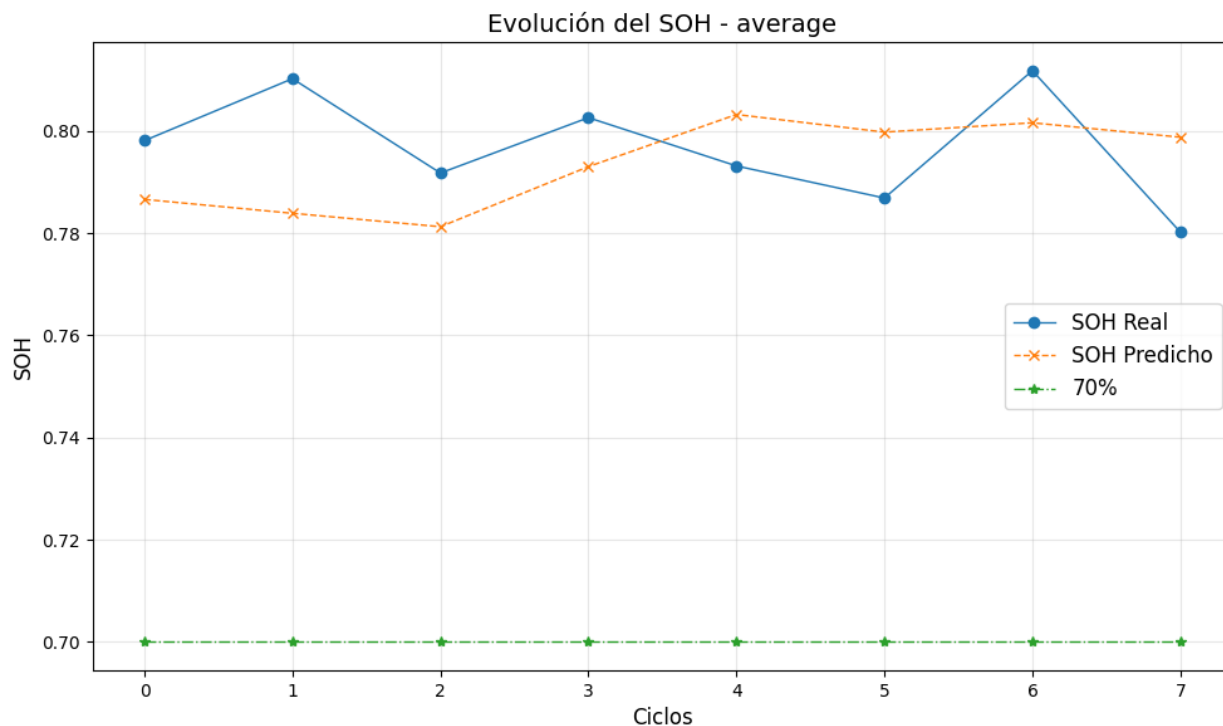


Figura 4.21: Predicciones del SoH de la batería en el vehículo intervenido utilizando el modelo MIT

Las predicciones de SoH de la batería en el futuro 10 ciclos en adelante, utilizando los datos de carga y descarga de la batería en el vehículo intervenido se pueden observar en la gráfica 4.21. En esta se denota claramente que el modelo acierta en gran medida los objetivos. También cabe destacar que el error obtenido en el modelo es comparable con las diferencias presentadas en la propia estimación del SoH target analizado en la sección 4.2.3. Por este motivo, a raíz de estos resultados se consideró cumplido el objetivo de la predicción de la degradación futura de la batería a partir de la predicción del valor de SoH.

4.3.9. Resumen

Tras evaluar el rendimiento de los modelos entrenados con ambos conjuntos de datos, se optó por implementar un enfoque de *ensamble* utilizando los 10 modelos obtenidos de la validación cruzada *k-fold* con el dataset del MIT. Esta estrategia permitió aprovechar las fortalezas individuales de cada modelo y mitigar sus debilidades específicas. Para la predicción

final, se ponderaron las salidas de cada modelo de forma equitativa. Este método de *ensemble* demostró una mejora significativa en la robustez de las predicciones, aumentando la precisión general del sistema, especialmente en condiciones operativas diversas, como lo es el caso de un vehículo en uso. La combinación ponderada de múltiples modelos permitió capturar mejor la complejidad de los patrones de degradación de las baterías y ofrecer estimaciones más confiables del estado de salud futuro, superando el poder de generalización de cualquier modelo individual.

4.3.10. Limitaciones del método elegido

Para cumplir el objetivo de poder estimar el SoH de cualquier vehículo eléctrico a partir de sus condiciones de funcionamiento actual, se utilizó un enfoque de aprendizaje automático específicamente, redes neuronales profundas en su formato de LSTM. Este camino tiene como pre-condición contar con una amplia cantidad de datos históricos sobre los cuales entrenar el modelo, optimizando los pesos de sus redes neuronales, y de esta forma poder interpretar relaciones complejas entre los datos.

A pesar de que estos modelos tengan una gran capacidad de generalización, el hecho de haber entrenado con datos de laboratorio en vez de datos extraídos directamente de vehículos en funcionamiento implica una limitación del método elegido. A esto usualmente se lo conoce como *domain adaptation*, es decir, utilizar un modelo en inferencia sobre datos distintos de los cuales fue entrenado. Cabe destacar que este proceso se ha realizado exitosamente en múltiples industrias como en el campo de *computer vision* así como también en el de predicción de series temporales [72].

Habiendo analizado los tres conjuntos de datos en profundidad, NASA, MIT y del vehículo intervenido, se identifica que existen diferencias en el comportamiento de las señales de voltaje, corriente y temperatura.

En primer lugar, la corriente en los conjuntos de datos de laboratorio es constante a tramos en la mayor parte, dado que los experimentos realizados se definen de esa forma estandarizada. En carga, esto no implica un problema dado que en los vehículos eléctricos la carga también es controlada mediante la corriente y el voltaje, siendo esta constante a secciones. En contraste, los datos del vehículo en descarga presentan variaciones en la corriente producto de la demanda variable de potencia en el funcionamiento común del mismo. No sólo la corriente no es constante, sino que también cambia de signo debido al frenado regenerativo.

Para abordar esta limitación, se realizaron diferentes pruebas de procesamiento de los datos de corriente vehículo. En primer lugar, se implementó un filtrado para eliminar los eventos de frenado regenerativo (corrientes positivas en descarga), con el objetivo de evaluar si al parecerse más a los datos de entrenamiento, esto mejoraba la precisión de las predicciones. Adicionalmente, se utilizó el valor absoluto de la corriente en los períodos de descarga. Estas modificaciones permitieron evaluar el impacto de agregar filtros en el pre-procesamiento de datos en el momento de inferencia. Sin embargo, no se obtuvieron buenos resultados de RMSE en comparación con los datos originales, por lo que no se incluyeron en el análisis.

Esto implica una clara limitante en el modelo obtenido, que no va a haber estado expuesto en el entrenamiento a los efectos que tienen estos cambios de corriente abruptos sobre la batería.

Por otra parte, también existen diferencias en las gráficas de tensión, aunque más sutiles que las analizadas anteriormente. En carga, las tres gráficas presentan una tendencia común de aumento rápido del voltaje seguido de una estabilización al acercarse a porcentajes de carga más altos. En descarga, presentan una tendencia opuesta, donde la tensión disminuye más lentamente para SoCs altos que para porcentajes de carga inferiores. Aunque los datos de tensión del auto preservan esta tendencia en términos generales, también presentan variaciones instantáneas más pronunciadas mientras el auto está en funcionamiento. Este último punto no se visualiza en las gráficas de los conjuntos de datos de laboratorio probablemente debido a las condiciones de carga y descarga a corriente constante definidas para el experimento.

Por último, la profundidad de descarga en ambos conjuntos de laboratorio era del 100 % para cada ciclo. En el funcionamiento real del vehículo, no es posible definir intervalos donde la profundidad de descarga sea total, considerando que en su funcionamiento es más común que se alternen cargas y descargas sin llegar a ser completas.

En la literatura consultada [73], se destaca que el comportamiento real de un vehículo con alternancia entre cargas y descargas, denominado *dynamic cycling*, es favorable para la vida útil de la batería. Indica que los perfiles de descarga de laboratorio subestiman la vida útil de las baterías hasta en un 38 %. Esta diferencia es sumamente significativa si se convierte esta cifra a kilómetros o millas, siendo esta de hasta 195.000 millas de diferencia.

Por los motivos anteriormente mencionados, se entiende que haber entrenado el modelo con datos de laboratorio permitió interpretar tendencias generales de la variación del SoH en las baterías. El volumen de datos proveído por estos conjuntos de datos sin duda fue una

gran ventaja para poder llegar a hacer predicciones con errores porcentuales bajos, incluso sobre los datos extraídos del vehículo. Además, los distintos tipos de carga introducidos por el dataset de MIT proveen una perspectiva valiosa respecto a los efectos de la carga rápida, es decir a mayor potencia, de las baterías eléctricas que se transfiere de una forma análoga a su utilización en EVs.

Sin embargo, no se puede ignorar que la estrategia utilizada presenta severas limitaciones, especialmente en cuanto a la captura de tendencias de utilización que son propias de su aplicación para vehículos, como lo son los picos de potencia debido a la aceleración, las corrientes positivas debido al frenado regenerativo y los extensos períodos de inutilización.

Se discutirá en mayor profundidad el posible trabajo a futuro para mitigar las limitaciones del método elegido en la sección Trabajo Futuro.

4.4. Análisis de patrones de degradación

4.4.1. Análisis de correlación de *features*

El análisis de correlación es una herramienta fundamental para comprender las relaciones entre las diferentes variables de un conjunto de datos. Evaluar estas correlaciones permite identificar patrones de comportamiento, redundancia de información y relaciones ocultas que pueden influir en el rendimiento de los modelos predictivos. Para este proyecto, se emplearon tres métodos tradicionales de cálculo de correlaciones: Pearson, Spearman y Kendall, cada uno con enfoques y propiedades particulares que permiten capturar distintos tipos de relaciones entre las variables.

1. Coeficiente de correlación de Pearson:

El coeficiente de Pearson mide la relación lineal entre dos variables continuas. Su valor varía entre -1 y 1, donde:

- Un valor cercano a 1 indica una correlación positiva fuerte (cuando una variable aumenta, la otra también tiende a aumentar).
- Un valor cercano a -1 indica una correlación negativa fuerte (cuando una variable aumenta, la otra tiende a disminuir).
- Un valor cercano a 0 sugiere que no existe una relación lineal entre las variables.

Este coeficiente se calcula mediante la siguiente expresión:

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

Donde x_i e y_i son los valores de las variables de la muestra i y n es la cantidad de observaciones o datos. Este coeficiente, a diferencia de los presentados a continuación, solamente permite captar patrones lineales [74].

2. Coeficiente de correlación de Spearman:

El coeficiente de Spearman mide la relación monotónica entre dos variables. Evalúa si una variable tiende a aumentar o disminuir de manera ordenada en relación con otra, independientemente de que la relación sea lineal. Su valor también varía entre -1 y 1 y se calcula a partir de los rangos de las variables:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Donde d_i es la diferencia entre los rangos de cada par de observaciones y n es el número total de observaciones. El coeficiente de Spearman es útil para detectar relaciones no lineales y es menos sensible a valores atípicos que el coeficiente de Pearson [74].

3. Coeficiente de correlación de Kendall:

El coeficiente de Kendall evalúa la relación ordinal entre dos variables. Se basa en el número de concordancias y discordancias entre pares ordenados de observaciones. Su valor también varía entre -1 y 1 y se define mediante la fórmula:

$$\tau = \frac{C - D}{\frac{1}{2}n(n - 1)}$$

Donde C es el número de pares concordantes, D es el número de pares discordantes y n es el número total de observaciones. El coeficiente de Kendall es más robusto frente a datos con distribuciones no normales y relaciones no lineales, lo que lo hace adecuado para conjuntos de datos complejos [74].

El uso combinado de estos tres métodos permite obtener una visión más completa de las relaciones entre las variables, ya que cada coeficiente destaca diferentes tipos de patrones y estructuras de dependencia. Sin embargo, los métodos tradicionales de correlación presentan limitaciones cuando las relaciones son altamente complejas o están influenciadas por múltiples factores de manera simultánea. Para abordar esta limitación, se complementó este análisis con el uso de valores Shapley, que permiten interpretar la contribución individual de cada variable dentro de un modelo complejo, capturando interacciones no lineales y relaciones de alta dimensión.

4.4.2. Valores Shapley

Comprender por qué un modelo realiza una determinada predicción puede ser tan importante como la precisión de la propia predicción en muchas aplicaciones. Hoy en día son los modelos más complejos, como los de aprendizaje profundo o los ensamblados, los que suelen ofrecer la mayor exactitud en grandes volúmenes de datos, sin embargo, es muy difícil interpretar cómo estos generan sus predicciones. Esto genera un desafío entre lograr precisión y mantener la interpretabilidad.

Los valores Shapley constituyen un método para explicar cómo un modelo realiza sus predicciones en función de la importancia de sus variables de entrada. El método fue desarrollado en el campo de la teoría de juegos como una forma justa de dividir el resultado de un juego entre sus jugadores y, posteriormente, adaptado al ámbito del aprendizaje automático para distribuir la predicción de un modelo entre las contribuciones de cada *feature*. Su aplicación permitió entender y depurar modelos al proporcionar una medida clara de la importancia de cada variable en el resultado final [75].

Esta herramienta proviene de un campo de la rama de inteligencia artificial explicada, XAI por sus siglas en inglés. Al calcular los valores Shapley de cada *feature* sobre modelos de caja negra, se pueden identificar patrones ocultos en los datos, incluyendo relaciones no lineales que estos modelos aprendieron de manera efectiva. La interpretación de los valores Shapley permite comprender cómo el modelo realiza sus predicciones y, en consecuencia, descubrir patrones previamente desconocidos en el conjunto de datos.

A diferencia de los métodos tradicionales de *feature importance*, que asignan una relevancia general a cada variable en el modelo, los valores Shapley permiten explicar predicciones individuales al indicar cómo cada *feature* o parámetro de entrada al modelo contribuye a la estimación realizada. La importancia de cada variable no se considera de manera aislada, sino en función del contexto proporcionado por el resto de las variables.

La teoría detrás de los valores Shapley se sustenta en tres propiedades fundamentales, que se resumen a continuación, las cuales garantizan una asignación equitativa de la contribución de cada variable a la predicción [75]:

1. Precisión local: La suma de todos los valores Shapley, más la predicción promedio del modelo, da como resultado la predicción final.

2. Ausencia de información: Cualquier *feature* que no alteró la predicción del modelo obtiene un valor Shapley de cero.

3. Aditividad: Los valores Shapley son aditivos, lo que los hace particularmente útiles para el análisis de *ensembles* de modelos, donde se puede calcular la contribución general de cada *feature* como el promedio ponderado de los valores Shapley de cada modelo base.

Debido a que cada valor Shapley representa una contribución marginal esperada, su cálculo requiere evaluar el impacto de cada variable en todas las combinaciones posibles de

valores de *features*, lo que implica un alto costo computacional en modelos con gran cantidad de variables.

La librería *open source* de Python SHAP, *SHapley Additive exPlanations*, busca resolver estas complejidades introduciendo métodos agnósticos al tipo de modelo, como Kernel SHAP, donde se utilizan aproximaciones a los valores shap para reducir el costo computacional que conlleva su cálculo.

4.4.2.1 Limitaciones del análisis mediante valores Shapley

El análisis mediante valores Shapley (SHAP) es una herramienta poderosa para interpretar modelos de *Machine Learning*, ya que permite identificar la contribución individual de cada característica en la predicción de un modelo. Sin embargo, presenta ciertas limitaciones que afectan su aplicabilidad y precisión en algunos casos específicos.

Kernel SHAP, a pesar de ser teóricamente agnóstico al modelo, no siempre lo es en la práctica, ya que no está implementado para todos los *frameworks* de aprendizaje automático. Aunque es compatible con las librerías de *Machine Learning* más populares, su aplicación en otros entornos puede resultar compleja y poco fiable. En el caso de Deep Explainer, diseñado para modelos de redes neuronales, su funcionamiento puede verse afectado por la estructura interna del modelo y la complejidad de sus capas.

Un problema clave en SHAP es la dependencia entre características. Kernel SHAP asume que las características de entrada son independientes entre sí al permutarlas para calcular las contribuciones marginales. Sin embargo, en presencia de correlación entre las variables, esta suposición deja de ser válida y puede dar lugar a resultados poco realistas o incluso contradictorios. Por ejemplo, si dos variables están altamente correlacionadas, SHAP puede asignar toda la contribución a una de ellas o repartirla de manera desigual, distorsionando la interpretación del modelo. Además, la permutación de valores correlacionados puede generar combinaciones de datos que no reflejan casos reales, lo que afecta la validez de las predicciones.

Otra limitación importante de SHAP es que no permite inferencia causal. Los valores SHAP explican cómo las características contribuyen a la predicción del modelo, pero no identifican las causas reales de la predicción. Esto se debe a que los modelos de machine learning están diseñados para detectar correlaciones, no causalidad. Por lo tanto, una alta contribución de una variable según SHAP no implica que esa variable sea una causa directa del resultado.

El factor humano también juega un papel en las limitaciones de SHAP. Es fácil interpretar erróneamente las contribuciones de las características, especialmente cuando los resultados parecen confirmar hipótesis previas o cuando existe un sesgo de confirmación. Además, el análisis de SHAP puede verse influenciado por errores humanos o por la tendencia a buscar patrones donde no existen, generando narrativas erróneas basadas en correlaciones espurias.

Finalmente, las limitaciones computacionales también deben considerarse. Kernel SHAP es computacionalmente costoso, ya que el cálculo de las permutaciones y la regresión lineal para estimar los valores de Shapley requiere un número exponencial de combinaciones posibles de características. Esto hace que SHAP sea difícil de aplicar en conjuntos de datos con una gran cantidad de variables o con alta dimensionalidad. Aunque TreeSHAP es más eficiente, está restringido a modelos basados en árboles y no puede aplicarse a otros tipos de modelos [76].

Estas limitaciones resaltan la importancia de interpretar los resultados de SHAP con cautela, especialmente en modelos complejos o con alta correlación entre las características. La combinación de SHAP con otros métodos explicativos y la consideración del contexto del modelo pueden ayudar a mitigar algunas de estas limitaciones y mejorar la comprensión de las predicciones.

4.4.3. Resultados de análisis de SHAP

Para este análisis de SHAP se realizaron dos tipos de gráficas que permiten visualizar diferentes aspectos de un parámetro de entrada a un modelo. Dadas las limitaciones vistas en la sección anterior, no se analizará el parámetro de entrada SoH_{actual} debido a que, aunque es útil para que el modelo realice las predicciones correctamente, no aporta información adicional sobre el funcionamiento del vehículo, dado que contiene la misma información que la variable $target$ pero con un desplazamiento temporal.

4.4.3.1 *Global Bar Plot:*

Estas gráficas permiten observar de manera ordenada que parámetros o *features* tienen mayor peso en la predicción de datos del modelo [77]. Estas gráficas se pueden realizar tanto para modelos individuales o, en caso de presentarse varios modelos con iguales parámetros de entrada y predicción como se haría en un proceso de *cross-validation*, se pueden promediar estas gráficas para obtener conclusiones generales de todo el entrenamiento. En este caso se realiza un proceso de *cross-validation* en el que se tienen 10 modelos entrenados a partir de 10 grupos o *batches* del *dataset* entero de MIT, con 140 baterías totales para analizar. Es sencillo realizar la gráfica de barras del *ensamble* gracias a la propiedad de aditividad de los valores Shapley.

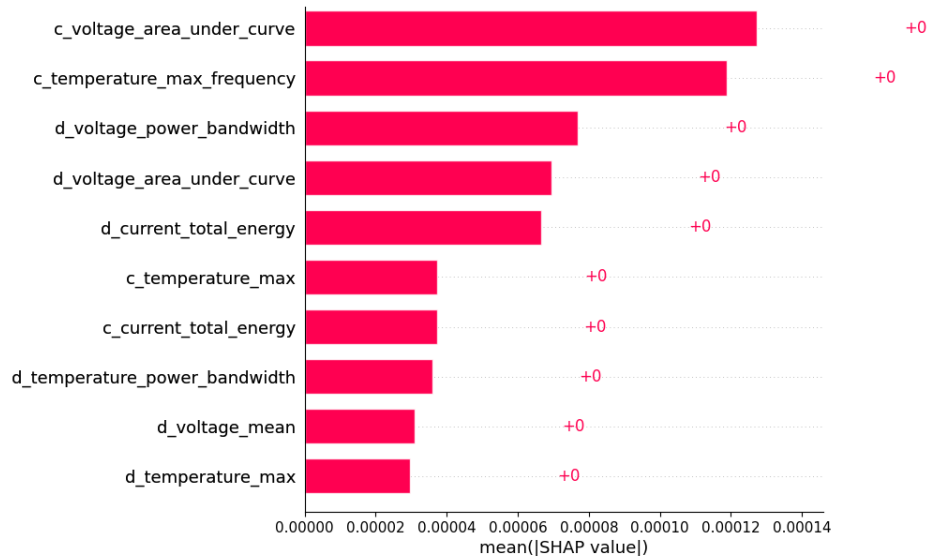


Figura 4.22: Gráfica de barras de *ensamble* de modelos.

De la figura 4.22, se obtiene el orden de los parámetros con mayor peso en la predicción. A mayor valor SHAP, mayor decisión tiene el parámetro en la predicción final. Los valores SHAP pueden ser tanto positivos como negativos, indicando si tienden a aumentar o a disminuir el valor de predicción final, respectivamente.

Los parámetros con prefijo "c_" hacen referencia a que fueron calculados a partir de datos de carga de las baterías, mientras que los parámetros con prefijo "d_" hacen referencia a que fueron calculados a partir de datos de descarga de la batería. Encabezando la lista de mayor peso en la predicción final se encuentran:

1. **c_voltage_area_under_curve:** El área debajo de la curva del voltaje en el proceso de carga.
2. **c_temperature_max_frequency:** La frecuencia de temperatura máxima calculada en una transformada rápida de Fourier, en el proceso de carga.
3. **d_voltage_power_bandwidth:** El ancho de banda de las frecuencias del voltaje obtenido a partir de la transformada rápida de Fourier, en el proceso de descarga.

De esta lista, la conclusión que se puede obtener es que aparentemente, el proceso de carga es más determinante de la degradación de la batería, recordando el entorno de laboratorio en el que estas fueron cargadas y descargadas. Esta gráfica no aporta más información que esta, con la siguiente gráfica se descubre nueva información.

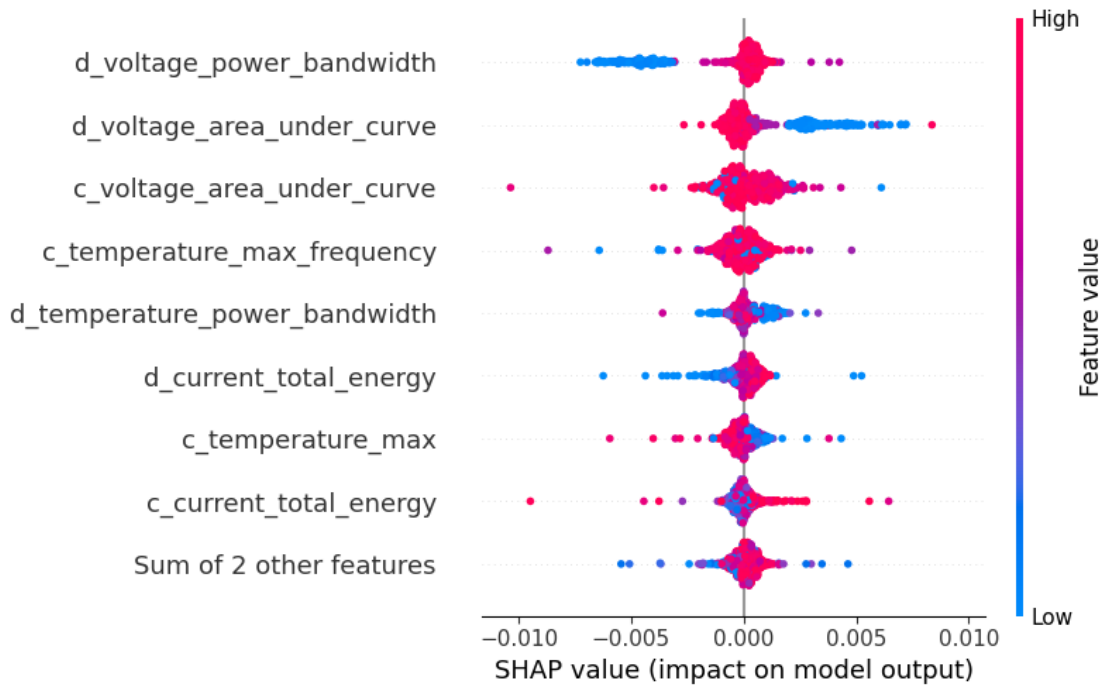
4.4.3.2 *Beeswarm Plot:*

Las gráficas tipo beeswarm permiten visualizar de manera compacta y detallada el impacto de las características más importantes en la salida de un modelo. Cada punto en el gráfico representa una instancia individual, donde la posición en el eje horizontal está determinada por el valor SHAP, que indica el efecto de esa característica en la predicción del modelo. La acumulación de puntos a lo largo de cada fila refleja la densidad de los valores SHAP, mientras que el color de cada punto representa el valor original de la característica correspondiente [78].

Una similitud con el gráfico de barras es que los parámetros a la izquierda se presentan en orden, siendo el primero el de mayor peso a la predicción.

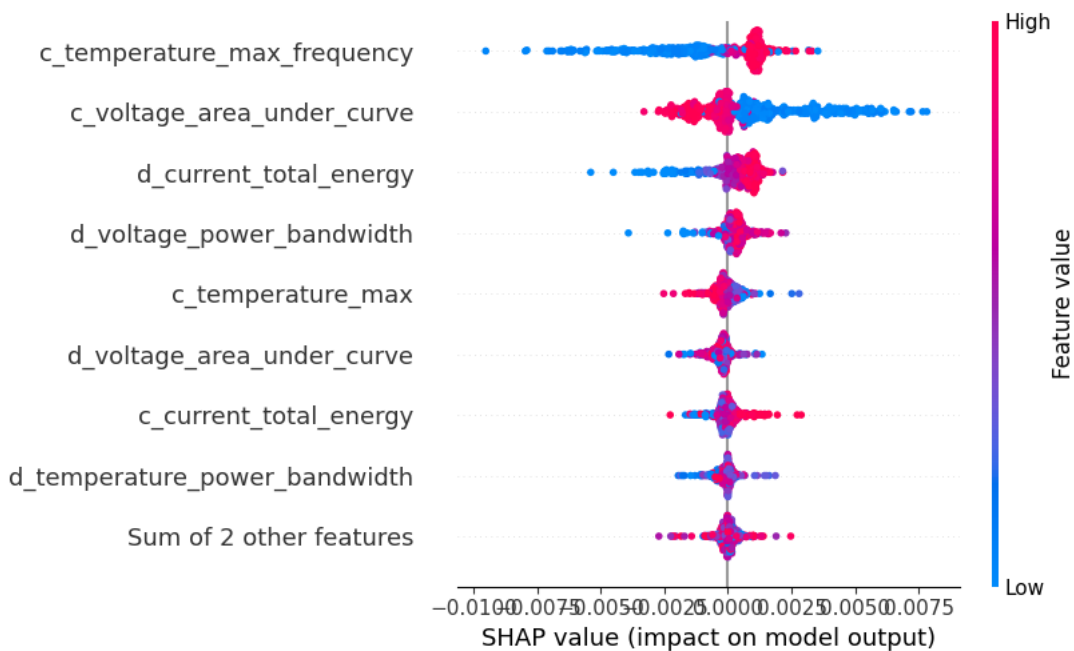
Sin embargo, a diferencia del gráfico anterior, estas gráficas no cumplen con la propiedad de aditividad ya que se representan las instancias individuales de los parámetros y promediar

estos valores no es correcto a nivel teórico. Es por esto que se presentan en la figura 4.23 dos gráficas de dos grupos o *fold*s individuales y sus resultados:



(a) *Fold 4*

Figura 4.23: Gráficas de *beeswarm* del análisis SHAP



(a) *Fold 6*

Estas gráficas permiten observar qué rango de valores de los parámetros de entradas tienden a aumentar o a disminuir la predicción del valor *target*. Por ejemplo, en el gráfico 4.24a (*Fold 6*), se realizan dos observaciones:

1. Para la variable "c.temperature_max_frequency" se observan puntos azules, indicando que valores bajos de esta variable, tienden a disminuir la predicción. En comparación, los puntos rosados, que representan los valores más altos de esta variable, parecen estar más agrupados cercano al valor SHAP cero (indicando que estos no tienen efecto considerable en la predicción).
2. Por el contrario, la variable "c.voltage_area_under_curve" parece indicar que valores bajos de este parámetro tienden a aumentar la predicción, en este caso, el estado de salud. Mientras tanto, una lectura análoga pasa para el rango superior de valores de esta variable.

Se elige la primera gráfica 4.23a (de nombre *Fold 4*) para observar que a pesar de utilizar los mismos datos de entrenamiento (rotando uno de los grupos de baterías para validación del modelo), ha entregado resultados completamente diferentes, indicando que las variables con mayor peso en la predicción del estado de salud de la batería son:

1. **d_voltage_power_bandwidth:** El ancho de banda de las frecuencias del voltaje obtenido a partir de la transformada rápida de Fourier, en el proceso de descarga.
2. **d_voltage_area_under_curve:** El área debajo de la curva del voltaje en el proceso de descarga.
3. **c_voltage_area_under_curve:** El área debajo de la curva del voltaje en el proceso de carga.

Parece ser que los parámetros calculados a partir de la *time series* de voltaje son los que tienen mayor peso en general, a través de todas las gráficas. El resto de los *beeswarm plots* se encontrarán en el anexo Gráficos *beeswarm* del análisis de SHAP.

4.4.3.3 Limitaciones del análisis

Desafortunadamente las conclusiones que se pueden obtener a partir de estas gráficas están limitada no sólo al método de análisis sino que también debido al modelo en sí mismo, particularmente en los datos de entrenamiento y prueba. Como ya se ha mencionado, los modelos han sido obtenidos a partir de diversos *datasets* donde las baterías son puestas a prueba en entornos de laboratorio, más controlados y expuestos en menor grado a efectos alternos que pueda generar el resto del vehículo. De esta limitante se habla en las secciones de Trabajo Futuro y Potencialidades del proyecto.

4.4.4. Conclusiones de ambos análisis

El objetivo final de este análisis es utilizar una combinación de los valores Shapley y las correlaciones de *features* para interpretar y desarrollar conclusiones sobre el estado de salud de las baterías y las posibles causas de su degradación. Para esto primero se realiza el análisis de SHAP sobre el modelo y los datos de los *datasets* para entender cómo el modelo hace sus predicciones. Luego, por separado y utilizando los datos obtenidos del Volkswagen e UP, se realizó un análisis de correlaciones que busca encontrar patrones de correlación entre los parámetros que el modelo utiliza para predecir y otros parámetros que pueden ser más fáciles de comprender y sobre los cuáles se tenga mayor control a la hora de operar el vehículo.

La lógica con la que se trabaja entonces es la siguiente. Llámese "X" a un *feature* o parámetro que afecta (según SHAP) a la predicción del SoH de alguna manera (ya sea positiva o negativa). Llámese "Y" a una variable que no es utilizada por el modelo para estimar el SoH (por ejemplo, la velocidad) y que se correlaciona fuertemente (ya sea positiva o negativamente) con "X". Pues entonces se infiere que ciertos valores o comportamientos de "Y" llevan a ciertos valores de "X" en una batería, lo que luego lleva a cierto impacto en el estado de salud de la misma. Que dos parámetros tengan correlación positiva implica que si una de las variables ("X" o "Y") aumenta, la otra también aumenta. Asimismo, si dos parámetros poseen correlación negativa entre ellos, cuando una aumente, la otra decrecerá.

A continuación se analizan algunos de los resultados de dicho análisis y sus posibles consecuencias.

Una relación que se puede observar con el análisis de correlaciones es la de "distancia total recorrida" (en un ciclo de descarga) y "Delta SoC" (es decir, la diferencia en el estado de carga al inicio y a final del intervalo de carga o descarga). En estos parámetros se ve una

correlación negativa fuerte, especialmente en el análisis de *Spearman* con un coeficiente de correlación de $-0,89$. Es decir a mayor distancia recorrida, más decrece el SoC. Esto es algo que intuitivamente parece lógico pero que no permite sacar conclusiones respecto al SoH. Sin embargo, es importante verificar que relaciones más triviales como esta sean reconocidas por el análisis.

Pasando a una correlación más interesante, se decidió buscar correlaciones con el *feature* "c_voltage_area_under_the_curve" (el área bajo la curva de voltaje durante la carga, mencionada ya en la sección 4.4.3) que, como se puede ver en las gráficas de *beeswarm*, mayores valores llevan a mayor degradación. La principal correlación que se encontró fue con la variable del SoC en la carga. En particular, en el análisis de Pearson, se ve un coeficiente de $0,93$ con el Delta SoC (SoC final menos el inicial) y también un coeficiente de $0,62$ con el SoC máximo del ciclo de carga (o sea el SoC final). Gracias a esto se puede concluir que cargas mas largas, y aún mas, cargas a mayores porcentajes generan un mayor valor de "c_voltage_area_under_the_curve" y por ende llevan a un mayor desgaste de la batería. Esto es consistente con recomendaciones populares de uso de autos eléctricos, donde normalmente se dice que es mejor operar a menores valores de SoC y es recomendado no cargar la batería al 100 % de no ser necesario.

Otro *feature* para la que se buscaron correlaciones fue la de "c_temperature_max_frequency", que a diferencia de la anterior genera mayor desgaste cuanto menor es su valor. En este caso se encontró una correlación que parece contradecir nuestra anterior conclusión, ya que a pesar de tener un impacto opuesto al *feature* anteriormente analizada, también se correlaciona positivamente con el Delta SoC, con un coeficiente de Spearman de 0.83 . Esto llevaría a pensar que cargar la batería a valores de SoC mayores es beneficioso, sin embargo es importante recordar las limitantes de estos métodos, del modelo, y de los *datasets* que se utilizan a la hora de analizar contradicciones como esta. Una mayor cantidad de datos reales de vehículos en diversas condiciones podría ayudar a aclarar el significado y la validez de esta contradicción. En cualquier caso, el impacto del *feature* de voltaje en la predicción de SoH es mayor que el de la frecuencia máxima de temperatura según sus valores SHAP, y su coeficiente de correlación con el Delta SoC también es mayor. Por lo tanto se concluye que; generalmente, evitar cargar la batería a estados de carga muy elevados es beneficioso para su salud a largo plazo.

5. Evaluación de la gestión de riesgos

En el desarrollo de este proyecto, se identificaron diversos riesgos técnicos y operativos que podrían afectar tanto la implementación como el funcionamiento del sistema de recopilación de datos en vehículos eléctricos. La identificación y gestión temprana de estos riesgos resulta fundamental para garantizar la estabilidad del sistema, la integridad de los vehículos involucrados y la precisión de los datos recopilados.

5.1. Falta de estandarización

El primer riesgo detectado está asociado a la diversidad en la arquitectura de comunicación de los vehículos eléctricos. La falta de un estándar universal para los parámetros de diagnóstico y la configuración de las tramas CAN implica que cada fabricante define su propia estructura de datos y comportamiento de las ECUs. Esta variabilidad entre marcas y modelos dificulta la tarea de decodificación de datos y podría limitar la capacidad del sistema para interpretar correctamente la información recopilada.

Gracias al apoyo de **Julio César Lestido**, quien brindó dos modelos de vehículos eléctricos, herramientas y sobre todo tiempo fue posible realizar este proyecto. No es menor la colaboración de **Smartway** que puso a disposición herramientas de *hardware* y de *software* para el análisis, así como el dispositivo que resultó instalado y permitió generar esta base de datos.

5.2. Acceso a vehículos eléctricos

Otro de los primeros riesgos considerados fue la posibilidad de no conseguir un vehículo eléctrico al cual conectarnos para realizar las pruebas de adquisición de datos. Sin embargo, este riesgo se mitigó desde las etapas iniciales del proyecto gracias al contacto temprano con varias marcas automotrices como Volkswagen, que demostraron apertura y disposición para colaborar. Esto permitió asegurar el acceso a vehículos eléctricos antes de que se iniciaran las etapas críticas del proyecto.

5.3. Interferencias en la comunicación CAN: posible causante de fallas

Se consideró especialmente el impacto que podría tener la instalación del dispositivo en la red de comunicación interna del vehículo, y la posibilidad de que el dispositivo a instalar obstruya la comunicación interna del mismo y pueda causar fallas en el funcionamiento y, en el peor de los casos, acabar con daños en los sistemas del vehículo o comprometer la seguridad del mismo.

Este riesgo fue mitigado al elegir utilizar el dispositivo de Smartway, un dispositivo confiable utilizado para monitorear más de un millón de kilómetros mensuales a través de diversos rubros [79]. Al inicio del proyecto se consideró la opción de diseñar un dispositivo propio *IoT* a efectos de realizar lo que el dispositivo SG de Smartway ya logra hace años. Se puede encontrar más información sobre este dispositivo en el anexo A.3.

Sin embargo, surge un problema durante el período de recopilación de datos que se explicará más adelante en la sección de Trabajo Futuro relacionado al funcionamiento definido en el vehículo y el enfoque no invasivo utilizado.

5.4. Incertidumbre en la cantidad de datos recopilados

Una fuente adicional de incertidumbre surgió en torno al tiempo efectivo de conexión al vehículo y si este lapso sería suficiente para recopilar una cantidad significativa de datos. En un principio, se estimó un período de dos meses como punto de partida, aunque no se contaba con una referencia clara de cuántos ciclos de carga y descarga podrían observarse en ese intervalo ni si los datos obtenidos serían suficientes para entrenar modelos predictivos. Como respuesta a esta limitación, y en busca de una mayor robustez en los modelos, se incorporaron al análisis conjuntos de datos públicos provenientes del MIT y la NASA. Estos datasets no formaban parte del plan original, pero surgieron como una solución alternativa ante la posible escasez de datos propios y permitieron realizar pruebas y entrenamientos más exhaustivos con volúmenes de datos representativos.

5.5. Resumen

La gestión de riesgos en este proyecto ha sido abordada mediante la implementación de dispositivos y protocolos diseñados para minimizar las interferencias con los sistemas del vehículo, optimizar el consumo energético y garantizar la continuidad de la transmisión de datos. El dispositivo SG de Smartway fue seleccionado específicamente por su compatibilidad con sistemas CAN y su probada estabilidad en entornos automotrices, lo que permite garantizar la fiabilidad y seguridad del sistema durante su operación.

6. Evolución del proyecto: Diagramas de Gantt

Desde la etapa inicial de planificación e investigación, el proyecto ha experimentado avances significativos que han permitido consolidar las bases técnicas y operativas para la implementación del sistema de monitoreo y análisis de datos en vehículos eléctricos.

1. Investigación y definición de objetivos:

La primera fase del proyecto se centró en la investigación y definición de los objetivos técnicos y operativos del sistema de monitoreo para vehículos eléctricos. Se comenzó con una exploración de las necesidades de los usuarios finales y las características de los sistemas eléctricos de los vehículos seleccionados. Este análisis permitió establecer los parámetros clave a monitorear, como la corriente de carga y descarga, voltaje, SOC (State of Charge), temperatura y consumo de energía. A partir de esta investigación inicial, se definieron los requisitos para el dispositivo *IoT*, las métricas de rendimiento y las condiciones de operación, asegurando que el sistema fuera compatible con múltiples modelos de vehículos y que pudiera adaptarse a diferentes protocolos de comunicación CAN. Esta fase fue crítica para alinear los objetivos técnicos con las capacidades reales del dispositivo y establecer un plan de acción claro para las siguientes etapas.

2. Investigación de normas y protocolos de comunicación:

Se llevó a cabo una investigación exhaustiva de las normas y protocolos de comunicación utilizados en la industria automotriz para la transmisión de datos de telemetría. La investigación se centró en protocolos como CAN (*Controller Area Network*), OBD-II (*On-Board Diagnostics*) y UDS (*Unified Diagnostic Services*), además de las normativas ISO y SAE que rigen el intercambio de información en vehículos eléctricos. Esta etapa

permitió comprender la estructura de las tramas CAN, las tasas de transmisión, los formatos de datos y los mecanismos de diagnóstico utilizados por los fabricantes.

3. Decodificación y análisis de datos:

La fase de decodificación se centró en la captura e interpretación de las tramas CAN transmitidas por el vehículo. Utilizando las herramientas *VehicleSpy* y ODIS, se logró identificar los identificadores de las tramas y los parámetros asociados a cada señal transmitida, como corriente, voltaje, temperatura y SOC. Se realizaron pruebas experimentales para validar la correspondencia entre las señales capturadas y las condiciones reales de operación del vehículo, correlacionando los datos obtenidos con la información mostrada en el tablero del automóvil. El proceso incluyó la identificación de tramas *multi-frame* y señales codificadas mediante técnicas propietarias del fabricante. Esta etapa permitió establecer una base sólida para el posterior análisis de datos, asegurando que las señales capturadas fueran precisas y representativas de las condiciones operativas del vehículo.

4. Acondicionamiento e instalación del dispositivo *IoT*:

Una vez completada la decodificación de las señales CAN, se procedió a la configuración y acondicionamiento del dispositivo *IoT* desarrollado por Smartway. El dispositivo fue programado para interpretar los parámetros CAN según las especificaciones obtenidas en la fase de decodificación. Se realizaron pruebas de laboratorio para validar la estabilidad y precisión de la lectura de datos, asegurando que el dispositivo pudiera operar de manera continua. Posteriormente, el dispositivo fue instalado en el vehículo, donde se realizaron pruebas de conexión y transmisión de datos en tiempo real. El sistema fue configurado para enviar datos de telemetría a un *endpoint* de Smartway mediante redes 4G y 2G, asegurando una transmisión estable y continua incluso en condiciones de baja cobertura.

5. Análisis de datos y elaboración de conclusiones:

Una vez instalado el dispositivo *IoT* y recopilados los datos de telemetría a través de la plataforma de Smartway, se inició el análisis de las señales obtenidas mediante el bus CAN del vehículo, como corriente, voltaje, temperatura, SOC (State of Charge) y estado del vehículo. Para facilitar el análisis, se implementó un proceso de pre-procesamiento de datos mediante un algoritmo de máquina de estados que permitió segmentar las series temporales en ciclos de carga y descarga. Esta segmentación permitió evaluar el comportamiento de la batería en diferentes escenarios operativos, como aceleración, fre-

nado regenerativo y tiempos de reposo, identificando patrones claros en las variaciones de corriente y voltaje.

6. Documentación y elaboración del informe final:

En la fase final, se consolidó toda la información obtenida a lo largo del proyecto en un documento técnico final. Este documento incluyó la descripción detallada del proceso de investigación, las técnicas de decodificación empleadas y el análisis de los datos recopilados. Se elaboraron gráficas y modelos para ilustrar la evolución del estado de salud de la batería y la dinámica de carga y descarga en función de las condiciones de operación. También se incluyeron recomendaciones técnicas para la mejora del sistema de monitoreo, así como propuestas para la optimización de la capacidad de la batería y la eficiencia energética del vehículo. El informe final sirvió como base para futuras implementaciones y permitió validar la efectividad del dispositivo *IoT* en condiciones reales de operación.

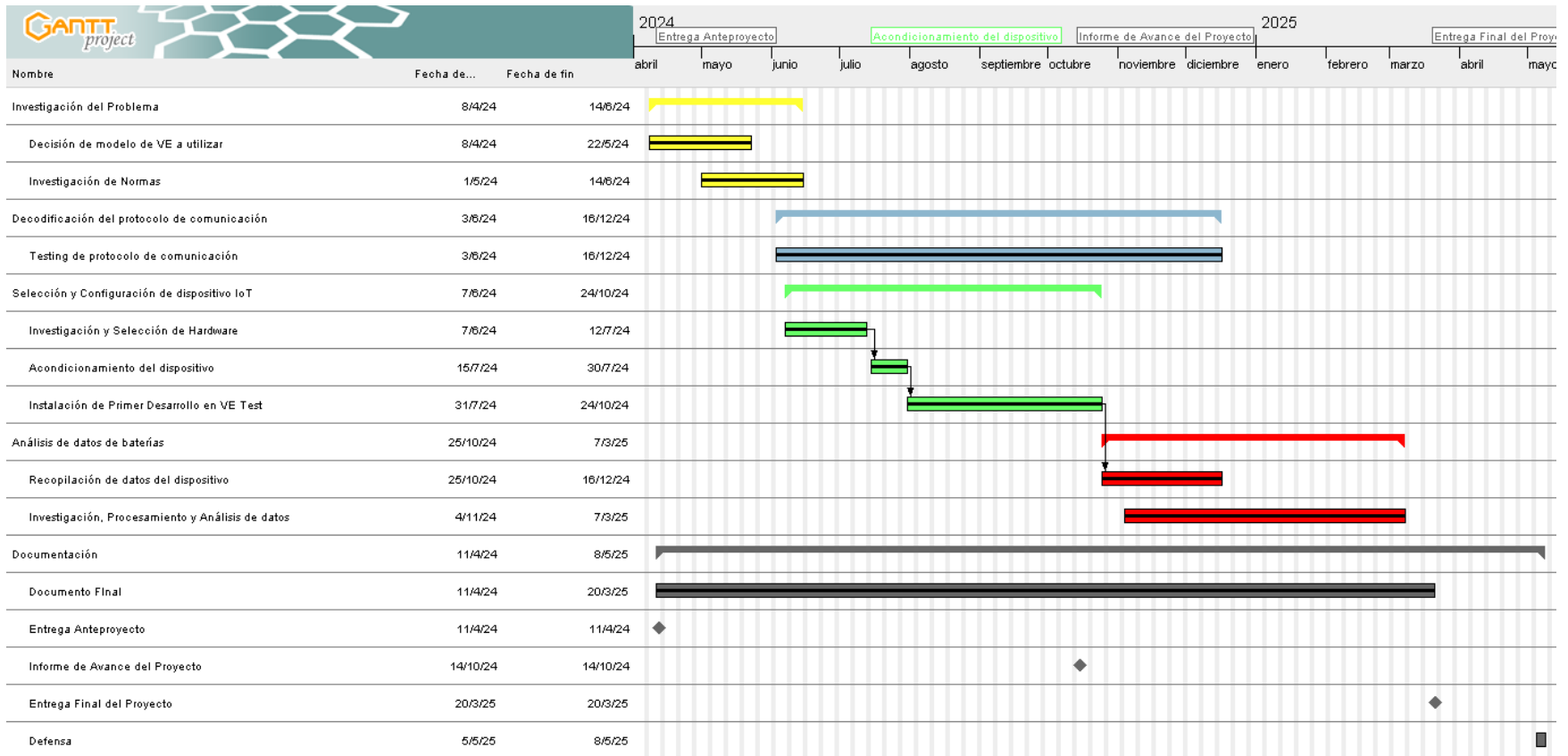


Figura 6.1: Diagrama de Gantt Final

7. Análisis Final

En esta etapa final del análisis se pueden resumir los logros más relevantes y constatar el cumplimiento de los objetivos trazados, así como la pertinencia de los resultados obtenidos. En la sección Objetivo, se planteó como propósito fundamental el análisis del estado de salud de las baterías en vehículos eléctricos y la identificación de los factores que inciden en su degradación. Para lograrlo, se establecieron una serie de Etapas del proyecto y Resultados clave, orientadas primero a la obtención de datos reales de los vehículos y, posteriormente, a su análisis mediante modelos predictivos de *Machine Learning*.

7.1. Cumplimiento de objetivos

ETAPA-1	Obtención de datos de vehículos.	Resultado esperado	Resultado obtenido
KR-1	Investigar y comprender en profundidad el protocolo de comunicación CAN, el sistema OBD-II, las ECU, el sistema BMS y todo material adicional que resulte relevante para el proyecto.	Hecho	Hecho
KR-2	Decodificar parámetros de dos vehículos eléctricos diferentes. Esto implica localizar los parámetros de interés para el análisis dentro de las unidades de control de los vehículos e interpretarlos. Una vez conocida la estructura para pedir estos datos, interpretar la respuesta y registrarla en una base de datos.	Hecho	Hecho*
KR-3	Reportar los datos obtenidos a una base de datos externa, por un período mínimo de 1 mes, para comenzar con la segunda etapa del proyecto.	Hecho	Hecho

Cuadro 7.1: Resultados de los objetivos de la etapa 1.

En cuanto al KR-2, consideramos el resultado obtenido dado que se pudo efectivamente decodificar todos los parámetros disponibles y relevantes en el bus can tanto del vehículo e-UP como del camión E-Delivery. Sin embargo, los parámetros necesarios para analizar el estado de la batería no estaban disponibles en el CAN del camión. Esto conllevó a que no se pudiera registrar su histórico, considerando que eso implicaría la instalación de sensores externos al vehículo que irían en contra de la premisa de la tesis de no realizar medidas invasivas en el vehículo.

ETAPA-2	Análisis de datos	Resultado esperado	Resultado obtenido
KR-1	Analizar los datos recopilados, conocer sus distribuciones y realizar un primer cálculo o estimación del estado de salud actual	Hecho	Hecho
KR-2	Investigación del estado del arte de las arquitecturas de Machine Learning en cuanto a la predicción de series temporales. Una vez realizado esto, entrenar y evaluar los modelos predictivos obtenidos, comparándolos con métricas relevantes.	Hecho	Hecho
KR-3	Obtener un análisis en base a los datos disponibles sobre qué aspectos afectan mayoritariamente a la degradación de la batería en el EV. El enfoque de el mismo será desarrollar estrategias para mitigar o ralentizar la degradación.	Hecho	Hecho

Cuadro 7.2: Resultados de los objetivos de la etapa 2.

7.2. Conclusiones

Uno de los principales hitos alcanzados fue la decodificación de los parámetros críticos de la red CAN en uno de los vehículos eléctricos. Este paso resultó vital para comprender la estructura de la información que provee el sistema de gestión de la batería (BMS). La instalación y prueba del dispositivo IoT permitió la recolección de datos de forma sistemática, representando así el cumplimiento de la primera meta del proyecto: disponer de información real del funcionamiento de los vehículos.

Una vez garantizada la toma de datos, la segunda etapa se centró en el tratamiento y la exploración de estas mediciones. El procesamiento y depuración de la información permitió identificar las variables con mayor correlación respecto al envejecimiento de la batería. Por ejemplo, el área bajo de la curva del voltaje tanto en carga como en descarga, la frecuencia máxima de la temperatura, la energía total de la corriente en descarga demostraron ser factores determinantes en la evolución del SoH.

El análisis posterior mediante modelos predictivos brindó evidencia de que es viable estimar el comportamiento futuro del SoH, aún cuando el conjunto de datos reales sea limitado y se deba combinar con bases de datos complementarias (NASA, MIT). De esta manera, se validaron los métodos de predicción para que, a mediano plazo, puedan aplicarse a más vehículos y en distintos entornos de operación.

A lo largo de la primera etapa, se logró la investigación y familiarización con los protocolos CAN y UDS, así como la comprensión de los principales componentes electrónicos (ECU y BMS), lo que permitió la decodificación de parámetros en dos vehículos eléctricos distintos. Esto culminó con la construcción y reporte de una base de datos, requisito indispensable para iniciar la segunda etapa, centrada en el análisis y modelado de la información obtenida. En este sentido, se realizó el examen exploratorio y el primer cálculo del estado de salud, seguido del entrenamiento y evaluación de diversos modelos predictivos. Finalmente, el estudio de la degradación permitió identificar los factores que ejercen mayor degradación e influencia en la evolución de la batería, a través del análisis de correlación y SHAP (XAI). A partir de este estudio, fue posible alcanzar diversas conjeturas para minimizar su deterioro, cumpliendo así con el propósito general de este proyecto de analizar y predecir el estado de salud de baterías de vehículos eléctricos en función de sus características de uso.

Si bien se cumplieron los objetivos principales, el proyecto enfrentó varias dificultades, entre ellas la escasa estandarización en la comunicación de distintos fabricantes y la necesidad de contar con volúmenes de datos representativos para afinar los modelos de predicción. Además, la extrapolación a una mayor variedad de vehículos y celdas con composiciones químicas diversas requerirá un mayor trabajo de campo y de verificación empírica.

7.3. Trabajo Futuro

El presente trabajo ha permitido establecer una base sólida para la recopilación y análisis de datos de vehículos eléctricos mediante un dispositivo de telemetría no invasivo. Sin embargo, se identificaron ciertas limitaciones y oportunidades de mejora que podrían ser abordadas en futuros desarrollos para optimizar el funcionamiento y la precisión de la recopilación de datos.

7.3.1. Volkswagen e UP

En el caso del auto eléctrico, se logra decodificar correctamente los parámetros determinantes para el entrenamiento de modelos de predicción del estado de salud de la batería y otros que permitirían adjudicar causas de degradación al perfil de conducción de los choferes o a los hábitos de carga del vehículo. Dado el método no invasivo utilizado, para conocer el cambio de estado de "apagado" a "moderando" o de "apagado" a "trabajando" (cargando en este caso), el dispositivo de recopilación de datos lo hacía despertando las ECUs del vehículo, provocando una descarga gradual en la batería. Esto implicó disminución en el estado de carga y una posible afectación al estado de salud de la batería. Se decidió el método menos invasivo posible debido a que el vehículo a analizar es parte de la flota de vehículos de alquiler de Julio César Lestido.

Al consultar con Smartway esta situación, se planteó una posible solución. Esta implica "invadir" al vehículo y encontrar alguna señal analógica dentro del vehículo (contacto o relé) para conocer el siguiente estado del vehículo cuando el mismo se encuentra apagado. El propio dispositivo sería capaz de detectar estas señales y retomar su funcionamiento normal cuando se conozca que el dispositivo está en conducción.

7.3.2. Volkswagen e Delivery

Respecto al camión eléctrico analizado, se podría continuar con su análisis, probando diferentes grados de "invasión" al vehículo.

El grado mínimo de invasión implicaría continuar analizando la herramienta ODIS que la sede de Volkswagen Camiones maneja en busca de los parámetros que se han encontrado en el vehículo anterior. Sin embargo, se tomaron capturas de todas las rutas posibles que ofrece el sistema ODIS y no se ha encontrado rastro de variables como voltaje, corriente y temperatura (claves para el desarrollo de modelos predictivos). Estos resultados fueron parte de las razones por las que se dejó de lado el análisis del segundo ejemplar eléctrico para enfocar esfuerzos en el análisis del primer modelo.

La experiencia de Smartway ha comentado que han analizado maquinaria pesada donde la comunicación CAN de diagnóstico no se ha presentado en el conector OBD. Ante esa situación, han buscado de manera interna la comunicación de diagnóstico y han encontrado pares torneados de cables. Con el objetivo de minimizar la interferencia electromagnética, el ruido inducido y el *crosstalk* es que las líneas de comunicación se presentan de esta manera. Esta es la segunda alternativa propuesta para completar el análisis al mismo nivel que su antecesor.

7.3.3. Análisis de datos y predicción de vida útil de la batería

La metodología aplicada en este proyecto para el análisis de la degradación de baterías se alinea con el estado del arte en la materia. En particular, el enfoque basado en análisis de datos explorado en esta tesis ha demostrado resultados prometedores y acelerados en múltiples disciplinas científicas. Además, la incorporación de técnicas de *Explainable AI* en este análisis permite obtener información valiosa para ampliar el conocimiento en el área.

No obstante, como se explicó previamente en la sección 4.3.10, este enfoque depende en gran medida de la calidad y representatividad de los datos con los que se entrena y valida. Se identificó que una de las principales limitaciones del proyecto fue la escasez de datos representativos de baterías en condiciones de uso similares a las de un vehículo eléctrico.

Por un lado, los datos de entrenamiento disponibles provenían de pruebas de laboratorio que no reflejan fielmente el uso real de los vehículos eléctricos. En modelos de aprendizaje

profundo, como las redes LSTM, es fundamental contar con un volumen considerable de datos para poder inferir tendencias significativas. Al iniciar el análisis, los conjuntos de datos accesibles al público correspondían a los proporcionados por el MIT y la NASA. Sin embargo, en diciembre de 2024 [73], se publicó un nuevo *dataset* con perfiles de carga más variados, diseñado para simular el comportamiento real de un vehículo eléctrico en condiciones de laboratorio. Este conjunto de datos incluso incorpora secciones con corriente de signo opuesto para replicar las fases de frenado regenerativo, una de las principales carencias de los *datasets* previos.

El acceso a este tipo de datos, no solo abre las puertas para poder realizar predicciones más precisas, sino que también permite calcular *features* más representativas de las series temporales de tensión, corriente y temperatura de los vehículos eléctricos, que cobran sentido cuando las corrientes dejan de ser constantes. Algunos ejemplos de estas, extraídas del artículo [73], son:

- **Corriente Promedio:** Valor promedio de la corriente durante el protocolo de carga y descarga.
- **Varianza Normalizada de la Corriente:** Variabilidad de la corriente a lo largo del protocolo, normalizada respecto a su valor promedio.
- **Corriente Máxima de Carga:** Valor máximo de corriente registrado durante la fase de carga.
- **Corriente Máxima de Descarga:** Valor máximo de corriente registrado durante la fase de descarga.
- **Número de Ocurrencias de Carga:** Número de veces que la corriente cambia a modo de carga desde un estado de reposo o descarga, representando eventos de frenado regenerativo.
- **Fracción Relativa de Carga:** Proporción del tiempo total del protocolo en el que la corriente es positiva ($I > 0$), indicando el tiempo dedicado a la carga.
- **Fracción Relativa de Reposo:** Proporción del tiempo total del protocolo en el que la corriente es cero ($I = 0$), representando periodos de reposo de la batería.
- **Fracción de Reposo a Alto SOC:** Porcentaje del tiempo de reposo en el que el estado de carga (SOC) es superior a 60 %, lo que puede afectar la degradación de la batería.

- **Relación Carga-Descarga:** Relación entre el tiempo total de carga y el tiempo total de descarga a lo largo del protocolo.
- **SOC en Reposo:** Promedio ponderado del estado de carga (SOC) durante los periodos de reposo de la batería, calculado en función de la capacidad.
- **Frecuencia Pico Intermedia:** Frecuencia más representativa dentro del espectro de señales asociado a las variaciones de corriente y voltaje durante el ciclo de carga y descarga.
- **Frecuencia Pico Máxima:** Frecuencia dominante en el espectro de señales de corriente y voltaje, identificada como la de mayor amplitud en el análisis de frecuencia.

La introducción de este conjunto de datos con el proyecto en su recta final, y considerando el tiempo que conllevó analizar en profundidad y pre-procesar los otros conjuntos de datos consultados, llevó a que no se pudiera incorporar para este estudio. Sin embargo, el hecho de que artículos académicos de esta índole estén siendo publicados resalta la vigencia de esta problemática en el rubro y la falta de datos representativos del funcionamiento de baterías en los vehículos eléctricos.

Por este motivo, se cree que aportaría mucho valor integrar este conjunto de datos en un análisis futuro.

Por otro lado, el volumen de datos que se pudo recopilar en el plazo del proyecto fue limitado. Con mayor tiempo de recolección de datos, hubiera sido posible eludir el mayor obstáculo del análisis: **el volumen de datos y la calidad de los datos.**

En el plazo de dos meses definidos para la extracción de datos se obtuvieron 46 procesos, 23 procesos de carga y 23 de descarga. Estos procesos de carga y descarga no se ajustan por completo a la definición de "ciclo" dada en la sección de Ciclo de carga y degradación de la batería. Con un mayor periodo de recolección, habría sido factible aplicar técnicas de *fine-tuning* o, en caso de contar con un volumen suficientemente grande, entrenar modelos de estimación del SoH exclusivamente con los datos recopilados.

Asimismo, disponer de datos de múltiples vehículos con características diferenciadas—ya sea en términos de fabricante, tamaño (autos, camiones, etc.) o frecuencia de uso (vehículos personales, empresariales, etc.)—sería de gran utilidad para validar los modelos predictivos desarrollados.

Con un modelo entrenado a partir de datos más representativos de escenarios reales, se podría mejorar su capacidad de generalización, logrando predicciones más precisas. Un modelo más fiel al comportamiento real, además, permitiría—mediante un análisis de SHAP—identificar con mayor precisión las relaciones funcionales entre las variables secundarias. Esto posibilitaría ajustar las lecturas e interpretaciones de las gráficas en función de los comportamientos específicos del vehículo tanto en carga como en descarga.

En resumen, se cree que sería de gran valor integrar al análisis los nuevos conjuntos de datos disponibles, así como también explorar la posibilidad de extraer una mayor variedad de datos de vehículos eléctricos en funcionamiento.

7.4. Potencialidades del proyecto

Este trabajo ha demostrado la viabilidad de implementar un sistema de monitoreo y análisis de datos para evaluar el estado de salud (SoH) de las baterías de vehículos eléctricos mediante telemetría no invasiva. Sin embargo, considerando un escenario con recursos ilimitados, incluyendo acceso a datos ilimitados, capacidad de procesamiento computacional y financiamiento sin restricciones, las potencialidades de esta tecnología podrían expandirse significativamente, abriendo nuevas oportunidades para optimizar el rendimiento y la vida útil de las baterías de manera más precisa y personalizada.

Una de las principales mejoras posibles en este contexto sería la posibilidad de recopilar y procesar una cantidad mucho mayor de datos de diversos vehículos eléctricos, cubriendo múltiples marcas y modelos bajo diferentes condiciones operativas y ambientales. La disponibilidad de una base de datos masiva permitiría entrenar modelos de predicción más complejos y robustos, capaces de capturar patrones de degradación específicos para cada tipo de batería y estilo de conducción. Esta capacidad de personalización permitiría generar recomendaciones específicas para el conductor o el operador de flota, optimizando el uso de la batería y extendiendo su vida útil. Incluso, permitiría justificar empíricamente los comentarios y sugerencias que se encuentran en internet sobre el buen cuidado de un vehículo eléctrico.

Con mayores recursos también se podría desarrollar una plataforma web en la que los usuarios de los vehículos monitoreados puedan obtener recomendaciones para ralentizar al máximo el estado de salud de la batería o incluso generar alertas en función de acciones que atenten más gravemente contra el estado de salud de la misma. Se discutía que esta plataforma podría generar reportes periódicos (mensual, trimestral, semestral, etc) para tener

registro del desgaste de la batería, detallando las mayores causas de degradación. Esto podría concientizar a los usuarios o choferes sobre el uso indebido de sus vehículos y cómo maximizar su vida útil. A esta idea se puede agregar un análisis de factores económicos para evaluar en términos monetarios el desgaste de la batería y cómo esta puede afectar al valor de reventa del vehículo.

Otra potencialidad sería la creación de un sistema de retroalimentación que, mediante actualizaciones OTA (*Over-The-Air*), permita mejorar continuamente los modelos de predicción y las estrategias de carga basadas en los datos de uso reales. Si se lograra un acceso continuo a los datos de los vehículos en tiempo real y una infraestructura computacional robusta, el modelo podría adaptarse dinámicamente a cambios en el rendimiento de la batería y en las condiciones de conducción, ajustando automáticamente las recomendaciones y las alertas sobre posibles fallos o desviaciones en el comportamiento de la batería.

En resumen, con acceso ilimitado a datos, financiamiento y capacidad de procesamiento, el sistema podría evolucionar hacia una plataforma de análisis predictivo avanzado y de optimización en tiempo real, capaz de proporcionar beneficios económicos directos y mejoras significativas en la eficiencia y longevidad de las baterías de vehículos eléctricos.

8. Referencias Bibliográficas

- [1] U.S. Department of Energy, “¿cómo funcionan los autos totalmente eléctricos?” n.d., recuperado el 14 de febrero de 2025. [Online]. Available: <https://afdc.energy.gov/vehicles/how-do-all-electric-cars-work>
- [2] What is a battery management system (BMS)? – how it works | synopsis. [Online]. Available: <https://www.synopsys.com/glossary/what-is-a-battery-management-system.html>
- [3] Volkswagen, “Baterías en los coches eléctricos: todo lo que necesitas saber,” n.d., recuperado el 18 de febrero de 2025. [Online]. Available: <https://www.vw.com/es/newsroom/everything-electric/batteries.html>
- [4] Apple Inc., “¿por qué baterías de iones de litio?” n.d., recuperado el 22 de febrero de 2025. [Online]. Available: <https://www.apple.com/es/batteries/why-lithium-ion/>
- [5] Facultad de Ingeniería, Universidad de la República Uruguay, Politécnica, INSIA, “Vehículos híbridos & eléctricos 2019: Características de las baterías para vehículos eléctricos,” 2019, presentación visitada el 23 de julio de 2024.
- [6] L. Ungurean, G. Cârstoiu, M. V. Micea, and V. Groza, “Battery state of health estimation: a structured review of models, methods and commercial devices: Battery State of Health Estimation: a Structured Review,” *International Journal of Energy Research*, vol. 41, no. 2, pp. 151–181, Feb. 2016. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/er.3598>
- [7] TÜVatio, “Modos de carga de un vehículo eléctrico,” n.d., recuperado el 23 de febrero de 2025. [Online]. Available: <https://tuvatio.es/blog/modos-de-carga-vehiculo-electrico/>
- [8] WapCar, “¿qué es chademo y ccs? todo lo que necesitas saber sobre la carga de vehículos eléctricos,” n.d., recuperado el 23 de febrero de 2025. [Online]. Available: <https://www.wapcar.my/news/what-is-chademo-and-ccs-everything-you-need-to-know-about-charging-evs-35378>
- [9] UTE Movilidad, “Red de carga para vehículos eléctricos en uruguay,” n.d., recuperado el 24 de febrero de 2025. [Online]. Available: <https://movilidad.ute.com.uy/carga.html?tab=red-de-carga>

- [10] Mecánica en Acción, “Escáner automotriz y protocolo j1962: Obd1 y obd2,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://mecanicaenaccion.com/diagnostico/escaner-automotriz-protocolo-j1962-obd1-y-obd2/>
- [11] “Bus CAN,” page Version ID: 162426726. [Online]. Available: https://es.wikipedia.org/w/index.php?title=Bus_CAN&oldid=162426726
- [12] CAN bus explained - a simple intro [2025]. [Online]. Available: <https://www.csselectronics.com/pages/can-bus-simple-intro-tutorial>
- [13] Winstar, “CAN Bus Interface Communication,” 2025, recuperado el 5 de marzo de 2025. [Online]. Available: <https://www.winstar.com.tw/es/can-bus-interface-communication>
- [14] Embien technologies: Guide to ISO 15765-2 or CAN-TP or DoCAN protocol. [Online]. Available: <https://www.embien.com/demystifying-iso-15765-2-can-tp-docan-protocol>
- [15] UDS explained - a simple intro (unified diagnostic services). [Online]. Available: <https://www.csselectronics.com/pages/uds-protocol-tutorial-unified-diagnostic-services>
- [16] Geotab, “What is obd-ii?” n.d., recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.geotab.com/blog/obd-ii/>
- [17] Intrepid Control Systems, “Vcan4-2 user guide,” n.d., recuperado el 28 de febrero de 2025. [Online]. Available: https://cdn.intrepidcs.net/guides/vcan4-2/vcan4-2_ug.pdf
- [18] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*. GitHub, 2020. [Online]. Available: https://d2l.ai/chapter_recurrent-modern/lstm.html
- [19] International Energy Agency. (2023) Electric vehicles. Accedido el 9 de abril de 2024. [Online]. Available: <https://www.iea.org/energy-system/transport/electric-vehicles>
- [20] E. P. Uruguay, “Movilidad sostenible: la polémica de las baterías de vehículos eléctricos y los avances científicos en uruguay,” 2023, recuperado el 17 de mayo de 2025. [Online]. Available: <https://www.elpais.com.uy/sostenible/movilidad-sostenible-la-polemica-de-las-baterias-de-vehiculos-electricos-y-los-avances-cientificos-en-u>
- [21] BloombergNEF, “Lithium-ion battery pack prices hit record low of \$139/kwh,” 2023, recuperado el 17 de mayo de 2025. [Online]. Available: <https://about.bnef.com/blog/lithium-ion-battery-pack-prices-hit-record-low-of-139-kwh>
- [22] P. Industry, “Rechargeable battery application notes (ap5),” n.d., recuperado el 17 de mayo de 2025. [Online]. Available: <https://industrial.panasonic.com/ww/ss/technical/ap5>

- [23] Aptiv, “What is an onboard charger?” 2023, recuperado el 17 de mayo de 2025. [Online]. Available: <https://www.aptiv.com/en/insights/article/what-is-an-onboard-charger>
- [24] V. España, “Componentes de un coche eléctrico,” n.d., recuperado el 17 de mayo de 2025. [Online]. Available: <https://www.volkswagen.es/es/electricos-hibridos/electricos/glosario/componentes-coche-electrico.html>
- [25] MathWorks, “Battery management system (bms),” n.d., recuperado el 14 de febrero de 2025. [Online]. Available: <https://la.mathworks.com/discovery/battery-management-system.html>
- [26] Honda España, “¿cómo funcionan las baterías de los coches eléctricos?” n.d., recuperado el 18 de febrero de 2025. [Online]. Available: <https://www.honda.es/cars/electric/how-do-electric-car-batteries-work.html>
- [27] Iberdrola, “Baterías de ion-litio: Innovación en almacenamiento de energía,” n.d., recuperado el 22 de febrero de 2025. [Online]. Available: <https://www.iberdrola.com/innovacion/baterias-ion-litio>
- [28] Redway Battery, “Densidad de energía de las baterías: Tipos y fórmula,” n.d., recuperado el 22 de febrero de 2025. [Online]. Available: <https://www.redway-tech.com/es/energy-density-of-batteries-types-and-formula>
- [29] B. University, “Bu-205: Types of lithium-ion,” 2024, accedido el 18 de mayo de 2025. [Online]. Available: <https://batteryuniversity.com/article/bu-205-types-of-lithium-ion>
- [30] S&P Global, “The evolution of electric vehicle battery technology,” 2024, accedido el 18 de mayo de 2025. [Online]. Available: <https://www.spglobal.com/automotive-insights/en/blogs/2024/11/the-evolution-of-electric-vehicle-battery-technology>
- [31] Coches.com, “Baterías lfp (litio ferrofosfato): Todo lo que debes saber,” n.d., recuperado el 23 de febrero de 2025. [Online]. Available: <https://noticias.coches.com/consejos/baterias-lfp-litio-ferrofosfato/504022>
- [32] EV Hub, “Volkswagen id.4 x 2024,” n.d., recuperado el 23 de febrero de 2025. [Online]. Available: <https://www.evhub.cc/portfolio/items/volkswagen-id4-x-2024>
- [33] Shop4Tesla, “Diferencias en las baterías de tesla,” n.d., recuperado el 23 de febrero de 2025. [Online]. Available: <https://www.shop4tesla.com/en/pages/tesla-unterschiede-der-batterie>

- [34] R. A. Appana, F. E. Idrissi, P. Ramesh, M. Canova, C. Y. Kang, and K. Um, “Diagnosing and Decoupling the Degradation Mechanisms in Lithium Ion Cells: An Estimation Approach,” May 2024, arXiv:2405.10857 [math]. [Online]. Available: <http://arxiv.org/abs/2405.10857>
- [35] I. Jorge, T. Mesbahi, A. Samet, and R. Boné, “Time Series Feature extraction for Lithium-Ion batteries State-Of-Health prediction,” *Journal of Energy Storage*, vol. 59, p. 106436, Mar. 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2352152X22024252>
- [36] International Electrotechnical Commission (IEC), *IEC 61851-1: Electric vehicle conductive charging system - Part 1: General requirements*, International Electrotechnical Commission Std. IEC 61 851-1:2017, 2017.
- [37] Ministerio de Industria, Energía y Minería de Uruguay, “Uruguay incorporó la primera estación de carga ultrarrápida para vehículos eléctricos,” n.d., recuperado el 24 de febrero de 2025. [Online]. Available: <https://www.gub.uy/ministerio-industria-energia-mineria/comunicacion/noticias/uruguay-incorporo-primera-estacion-carga-ultrarrapida-para-vehiculos>
- [38] Volkswagen Canarias, “El freno regenerativo de un coche eléctrico,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://www.vwcanarias.com/es/blog/el-freno-regenerativo-de-un-coche-electrico.html>
- [39] Senthil Kumar P R, “A brief history of automotive ecu evolution,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://www.linkedin.com/pulse/brief-history-automotive-ecu-evolution-senthil-kumar-p-r-ec0jc/>
- [40] Altium, “Controller area network (can) bus: Introduction and history,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://resources.altium.com/p/Controller-Area-Network-Bus-Introduction-and-History>
- [41] GaragePro, “History of on-board diagnostics (obd),” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://garagepro.shop/blogs/all-things-obd/history-of-on-board-diagnostics>
- [42] CAN in Automation (CiA), “History of can technology,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://www.can-cia.org/can-knowledge/history-of-can-technology>

- [43] Vector eLearning, “Introduction to can technology,” n.d., recuperado el 25 de febrero de 2025. [Online]. Available: <https://elearning.vector.com/mod/page/view.php?id=335>
- [44] Organización Internacional de Normalización (ISO), *Vehículos de carretera — Diagnóstico en red de área de controladores (DoCAN) — Parte 2: Capa de transporte y capa de red*, ISO Std., 2024, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/84211.html>
- [45] —, *Vehículos de carretera — Servicios de diagnóstico unificados (UDS) — Parte 1: Especificación de la aplicación de servicios de diagnóstico*, ISO Std., 2006, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/45293.html>
- [46] Automotive Test Solutions, “Understanding the controller area network (can),” n.d., recuperado el 26 de febrero de 2025. [Online]. Available: <https://automotivetestsolutions.com/knowledge-base/case-studies/all-case-studies/understanding-the-controller-area-network-can>
- [47] Organización Internacional de Normalización, *Vehículos de carretera — Red de área de controladores (CAN) — Parte 2: Requisitos de la capa física de alta velocidad*, ISO Std., 2016, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/67244.html>
- [48] Organización Internacional de Normalización (ISO), *Vehículos de carretera — Red de área de controladores (CAN) — Parte 2: Requisitos de la capa física de alta velocidad*, ISO Std., 2024, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/85120.html>
- [49] —, *Vehículos de carretera — Servicios de diagnóstico unificados (UDS) — Parte 3: Servicios de transporte de datos (DTS)*, ISO Std., 2022, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/77323.html>
- [50] —, *Vehículos de carretera — Red de área de controladores (CAN) — Parte 1: Disposición de la capa de acceso al medio y la capa física dependiente del medio*, ISO Std., 2024, recuperado el 26 de febrero de 2025. [Online]. Available: <https://www.iso.org/standard/86384.html>
- [51] S. Sites. CAN bus voltage level: Analyzing the electrical potential. [Online]. Available: <https://sitaltech.com/can-bus-voltage-level-analyzing-the-electrical-potential/>
- [52] CAN FD. [Online]. Available: <https://www.can-cia.org/can-knowledge/can-fd-the-basic-idea>

- [53] OBDOnUDS - SAE j1979-2 - diagnostic standard. [Online]. Available: <https://www.vector.com/us/en/products/solutions/diagnostic-standards/obdonuds/>
- [54] ZEVonUDS - zero emission vehicle SAE j1979-3 diagnostic standard. [Online]. Available: <https://www.vector.com/us/en/products/solutions/diagnostic-standards/zevonuds-zero-emission-vehicle-sae-j1979-3-diagnostic-standard-vector/>
- [55] TechRoute66, “Odis-s: Diagnóstico para vehículos del grupo volkswagen,” n.d., recuperado el 26 de febrero de 2025. [Online]. Available: <https://techroute66.com/es/producto/odis-s/>
- [56] GoingElectric, “Liste der OBD2-codes,” n.d., recuperado el 4 de marzo de 2025. [Online]. Available: <https://www.goingelectric.de/wiki/Liste-der-OBD2-Codes/>
- [57] CSS Electronics, “J1939 Explained - A Simple Intro (2025 Tutorial),” 2025, recuperado el 5 de marzo de 2025. [Online]. Available: <https://www.csselectronics.com/pages/j1939-explained-simple-intro-tutorial>
- [58] AutoPi, “Autopi: The open automotive iot platform,” n.d., recuperado el 28 de febrero de 2025. [Online]. Available: <https://www.autopi.io/>
- [59] Universidad Católica del Uruguay (UCU), “Optimización de la vida Útil de las baterías en vehículos eléctricos,” n.d., recuperado el 28 de febrero de 2025. [Online]. Available: <https://liberi.ucu.edu.uy/xmlui/bitstream/handle/10895/1690/Optimizaci%C3%B3n%20de%20la%20Vida%20%C3%A9til%20de%20las%20Bater%C3%ADas%20en%20Veh%C3%ADculos%20El%C3%A9ctricos.pdf?sequence=1&isAllowed=y>
- [60] A. Gailani, R. Mokidm, M. El-Dalahmeh, M. El-Dalahmeh, and M. Al-Greer, “Analysis of lithium-ion battery cells degradation based on different manufacturers,” *arXiv preprint arXiv:2007.01937*, 2020. [Online]. Available: <https://arxiv.org/pdf/2007.01937>
- [61] I. Jorge, T. Mesbahi, A. Samet, and R. Boné, “Time Series Feature extraction for Lithium-Ion batteries State-Of-Health prediction,” *Journal of Energy Storage*, vol. 59, p. 106436, Mar. 2023.
- [62] NASA Ames Prognostics Center of Excellence, “Li-ion battery aging datasets,” <https://data.nasa.gov/dataset/Li-ion-Battery-Aging-Datasets/uj5r-zjdb>, recuperado en julio de 2025.
- [63] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, and R. D.

- Braatz, “Data-driven prediction of battery cycle life before capacity degradation,” *Nature Energy*, vol. 4, no. 5, pp. 383–391, May 2019.
- [64] P. Khumprom and N. Yodo, “A Data-Driven Predictive Prognostic Model for Lithium-ion Batteries based on a Deep Learning Algorithm,” *Energies*, vol. 12, no. 4, p. 660, Feb. 2019.
- [65] Z. Lv, Y. Song, C. He, and L. Xu, “Remaining useful life prediction for lithium-ion batteries incorporating spatio-temporal information,” *Journal of Energy Storage*, vol. 88, p. 111626, May 2024.
- [66] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [67] B. D. Fulcher and N. S. Jones, “Highly comparative feature-based time-series classification,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 12, pp. 3026–3037, 2014.
- [68] NVIDIA cuDNN — NVIDIA cuDNN. [Online]. Available: <https://docs.nvidia.com/deeplearning/cudnn/latest/>
- [69] Python release python 3.10.0. [Online]. Available: <https://www.python.org/downloads/release/python-3100/>
- [70] CUDA toolkit 11.2 downloads. [Online]. Available: <https://developer.nvidia.com/cuda-11.2.0-download-archive>
- [71] cuDNN archive. [Online]. Available: <https://developer.nvidia.com/rdp/cudnn-archive>
- [72] Z. Li, R. Cai, T. Z. J. Fu, Z. Hao, and K. Zhang, “Transferable time-series forecasting under causal conditional shift,” *arXiv preprint arXiv:2111.03422*, 2023, accepted in *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. [Online]. Available: <https://arxiv.org/abs/2111.03422>
- [73] A. Geslin, L. Xu, D. Ganapathi, K. Moy, W. C. Chueh, and S. Onori, “Dynamic cycling enhances battery lifetime,” *Nature Energy*, Dec. 2024. [Online]. Available: <https://www.nature.com/articles/s41560-024-01675-8>
- [74] Statistics Solutions, “Correlation: Pearson, kendall, spearman,” <https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/correlation-pearson-kendall-spearman/>, recuperado el 9 de marzo de 2025.

- [75] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.
- [76] Conor O’Sullivan, “The limitations of shap,” <https://medium.com/data-science/the-limitations-of-shap-703f34061d86>, 2022, recuperado en agosto de 2025.
- [77] SHAP Developers, “SHAP Bar Plot Example,” https://shap.readthedocs.io/en/latest/example_notebooks/api_examples/plots/bar.html, 2025, recuperado el 14 de marzo de 2025.
- [78] —, “SHAP Beeswarm Plot Example,” https://shap.readthedocs.io/en/latest/example_notebooks/api_examples/plots/beeswarm.html, 2025, recuperado el 14 de marzo de 2025.
- [79] SmartWay, “On Smart Way Official Website,” 2025, recuperado el 8 de marzo de 2025. [Online]. Available: <https://www.onsmartway.com/>

A. Anexos

A.1. Resultados de Volkswagen e UP (auto eléctrico)

ECU #	Nombre de ECU
01	Electrónica del motor
17	Cuadro de instrumentos
23	Servofreno
25	Inmovilizador
51	Accionamiento eléctrico
BD	Sistema de carga de batería de alto voltaje
C6	Cargador de baterías de alto voltaje

Cuadro A.1: Lista de ECUs y sus nombres en VW e UP

Parámetro	Conversión
Kilometraje	BB*256 + CC
Borrador de averías sin confirmar	AA

Cuadro A.2: Tabla de referencia para parámetros de kilometraje y averías

Código	Estado del Cargador
0	No cargando
1	Cargando AC
2	Reservado
3	Cargando CC
4	Precargar
5	Fallo interno
6	Reservado
7	Inicialización

Cuadro A.3: Tabla de referencia para los estados del cargador

Parámetro	Conversión	Comentario
Tensión efectiva de batería de alto voltaje	$0,25 * (AA*256 + BB)$	
Corriente efectiva de batería de alto voltaje	$0,25 * (CC*256 + DD)$	
Punto de ajuste climático / Solicitud de climatización	EE	<ul style="list-style-type: none"> * 0: Enfriamiento alto * 1: Enf. promedio * 2: Enf. bajo * 3: Sin solicitud * 4: Calor bajo * 5: Calor promedio * 6: Calor alto * 7: Circulación
Modo BMS	FF	<ul style="list-style-type: none"> * 0: En espera * 1: Enfriamiento bajo * 2: Equilibrio * 3: Carga externa mediante cargador * 4: Carga externa * 5: Error * 6: Carga de CC * 7: Inicialización
Límite de corriente de carga	$GG*256 + HH$	
Tensión de carga máxima para la batería	$II*256 + JJ$	
Temperatura de la batería HV (°C)	KK	

Cuadro A.4: Tabla de referencia para parámetros adicionales de la batería

Cuadro A.5: Tabla de parámetros CAN

Parámetro	Tx Header	Rx Header	ECU #	ID	Conversión	Unidad
VIN	\$711	\$77B	25	F1 90	-	17 dígitos alfanum
Keep Alive	\$700	-	-	3E 80	-	-
Battery %	\$765	\$7CF	BD	1D D0	AA / 2	%
<i>BMS values</i>	\$765	\$7CF	BD	1D E6	Tabla A.4	-
Velocidad	\$7E0	\$7E8	01	F4 0D	AA	km/h
Kilometraje cada 10km	\$714	\$77E	17	22 03	$(AA*256 + BB) * 10$	km
Capacidad	\$714	\$77E	17	22 E4	Este valor puede funcionar como estimación de SoH si $SoC = 100\%$ y $T_{bat} = 25^{\circ}C$	kWh

Parámetro	Tx Header	Rx Header	ECU #	ID	Conversión	Unidad
Consumo	\$714	\$77E	17	22 E3	$0,1 * (AA*256 + BB)$	kWh/100 km
Par Valor Real	\$7E6	\$7EE	51	3E E9	$\text{signed}(AA*256 + BB) / 16$ Complemento a 2	Nm
Datos ambientales 1	\$73B	\$7A5	23	02 BD	Tabla A.2	-
Tensión en borne +15	\$7E6	\$7EE	51	02 85	OFF = 0 ON = 1	CATEGÓRICA
Tensión en borne +30	\$7E0	\$7E8	01	F4 42	$0.001 * (AA*256+BB)$	V
Tensión en borne +50	\$7E0	\$7E8	01	12 6C	No llave a fondo = 0 Llave a fondo = 1	CATEGÓRICA
Cargador	\$744	\$7AE	C6	1D A0	Tabla A.3	CATEGÓRICA
<i>Power Efficiency</i>	\$744	\$7AE	C6	15 D6	$AA*0.1 + 75$ Eficiencia energética cuando el cargador se encuentra conectado	%

167

Continúa en la siguiente página...

Parámetro	Tx Header	Rx Header	ECU #	ID	Conversión	Unidad
<i>AC Input Current</i> (cargando)	\$744	\$7AE	C6	41 FB	Modulo 1: 0.1 * AA Modulo 2: 0.1 * BB Modulo 3: 0.1 * CC	A
<i>AC Input Voltage</i> (cargando)	\$744	\$7AE	C6	41 FC	Modulo 1: AA*256 + BB Modulo 2: CC*256 + DD Modulo 3: EE*256 + FF	V
Corriente de carga CC	\$765	\$7CF	BD	1D E2	Corriente MAX: AA Corriente Real: BB	A
Voltaje de carga CC	\$765	\$7CF	BD	1D E3	Voltaje MAX: AA*256+BB Voltaje Real: CC*256+DD	V
Tiempo restante de carga	\$765	\$7CF	BD	1D E4	AA*5	minutos

A.2. Conectores necesarios para la decodificación de Volkswagen e Delivery (camión eléctrico)

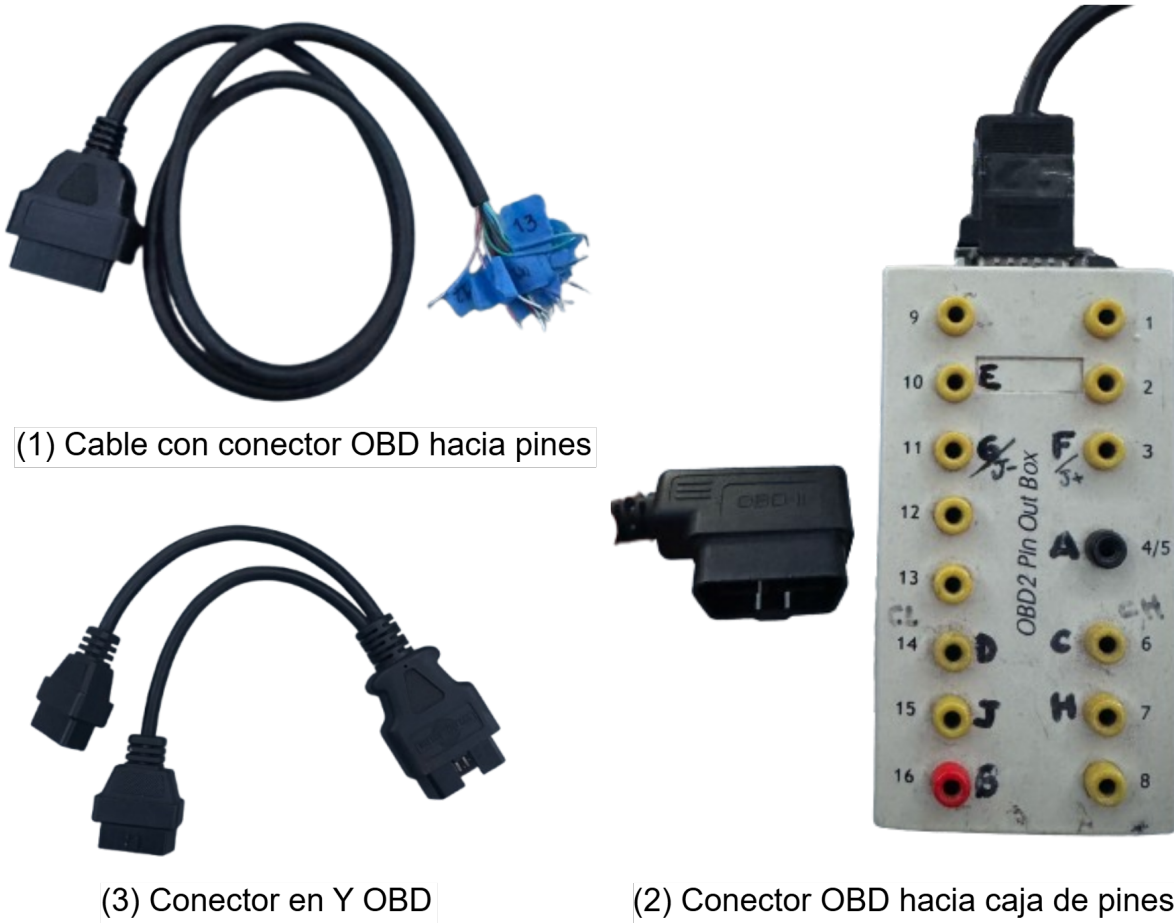


Figura A.1: Imágenes de conectores para decodificación del camión eléctrico

A.3. Dispositivo SG de Smartway: TDU

El dispositivo TDU (Telematic Diagnostic Unit) es un equipo de telemetría desarrollado íntegramente por Smartway, tanto a nivel de hardware como de software. Su función principal es recopilar y transmitir datos de diagnóstico de vehículos mediante la lectura de redes de comunicación internas como CAN y otros protocolos industriales estandarizados.

El TDU se conecta a las redes de datos internas del vehículo a través del puerto OBD-II o mediante conexiones CAN directas. La interpretación de la información recibida se realiza mediante una configuración específica que se descarga al dispositivo, lo que permite traducir y procesar parámetros como temperatura, revoluciones por minuto (RPM), velocidad, voltaje, corriente y estado de carga de la batería.

El dispositivo transmite estos datos periódicamente a una plataforma web mediante redes 2G o 4G, según la configuración establecida. Si la red celular no está disponible, el TDU almacena los datos internamente hasta por una semana y los reporta automáticamente cuando la conexión se restablece. Además, cuenta con un módulo de geolocalización (GPS) que permite asociar la información de diagnóstico con la posición del vehículo en tiempo real.

El TDU procesa y administra los datos obtenidos mediante las redes de comunicación del vehículo, estructurándolos en paquetes de datos que son posteriormente transmitidos a la plataforma web de monitoreo. El proceso de manejo de tramas incluye:

- Captura de datos: La información recibida por el bus CAN es analizada mediante la configuración cargada en el dispositivo, permitiendo identificar y procesar parámetros específicos.
- Almacenamiento y retención: Si el dispositivo pierde la conexión con la red celular, los datos son almacenados internamente en la memoria flash hasta por una semana, garantizando la continuidad del proceso de registro y análisis.
- Transmisión de datos: Los datos almacenados y los datos en tiempo real se transmiten mediante las redes 2G o 4G según la disponibilidad de la red y la configuración del dispositivo.
- Análisis y correlación: La información se correlaciona con parámetros de geolocalización y estado del vehículo para generar reportes detallados accesibles desde la plataforma web de Smartway.

LED	Estado	Descripción
Rojo (Power)	Apagado	Sin alimentación.
	Parpadeando rápido (0.25 s)	Alimentación desde batería interna.
	Encendido	Alimentación principal activa.
Rojo (Status)	Apagado	Vehículo apagado.
	Parpadeando rápido (0.25 s)	Vehículo en contacto.
	Parpadeando lento (1 s)	Vehículo moderando.
	Encendido	Vehículo en marcha.
Verde (GPS)	Apagado	GPS no funciona.
	Parpadeando rápido (0.25 s)	GPS sin señal.
	Encendido	GPS funcionando correctamente.
Verde (Celular)	Apagado	Módem sin registrar.
	Parpadeando rápido (0.25 s)	Registrando en la red.
	Parpadeando lento (1 s)	Registrado en roaming.
	Encendido	Módem registrado en la red.

Cuadro A.6: Indicadores luminosos del dispositivo TDU

A.4. Visualización de los datos MIT

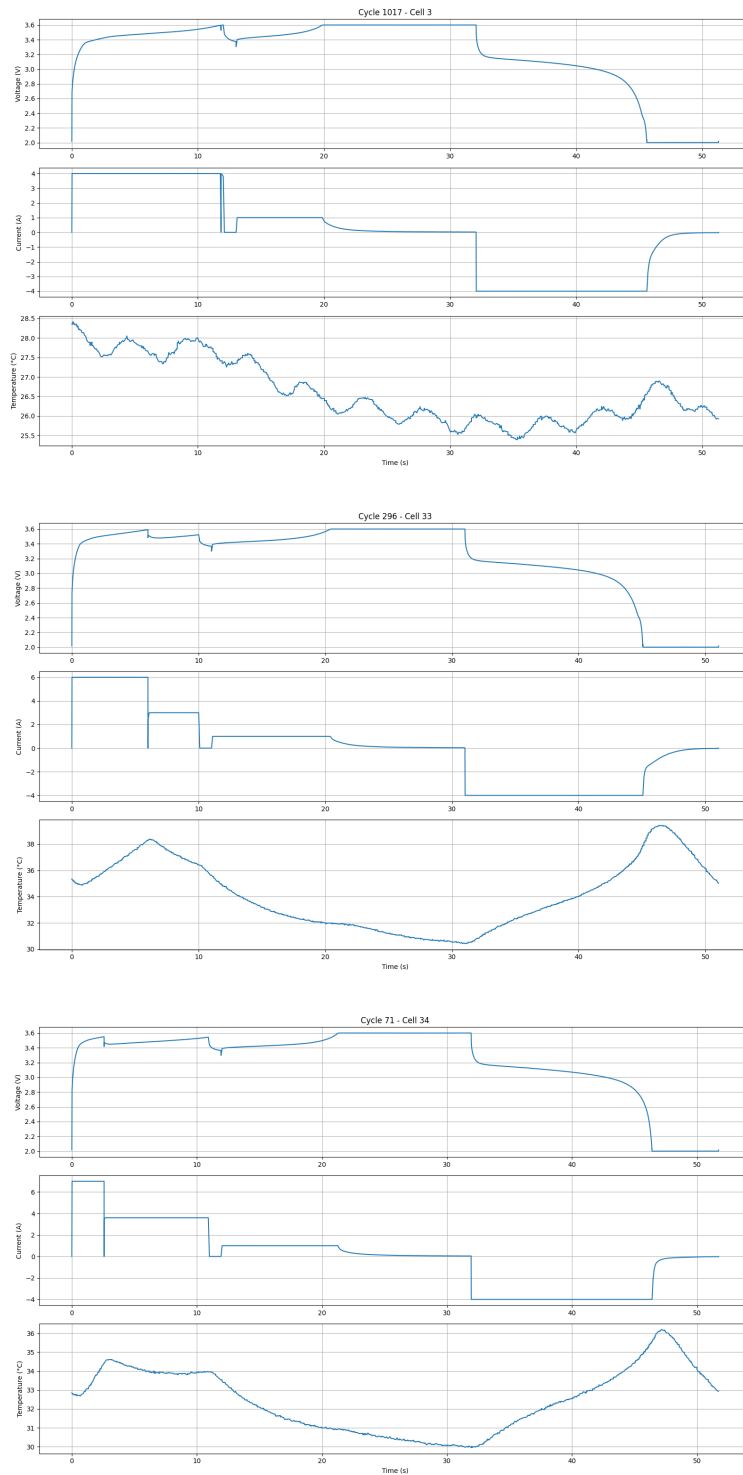
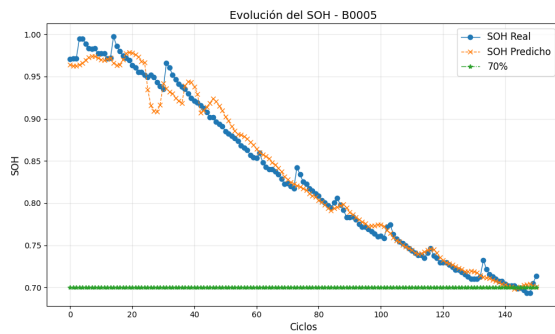
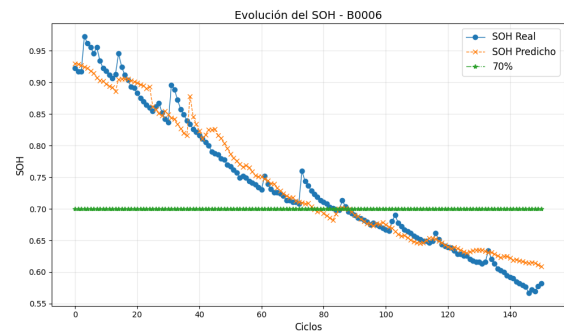


Figura A.2: Visualización de las señales en celdas MIT

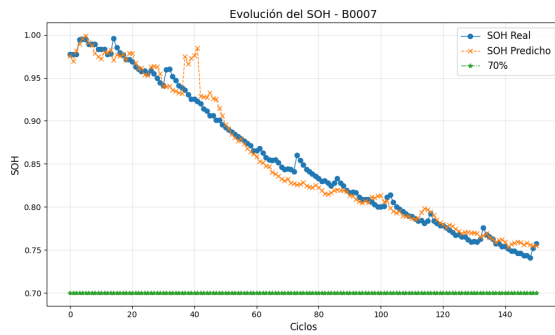
A.5. Resultados modelos NASA



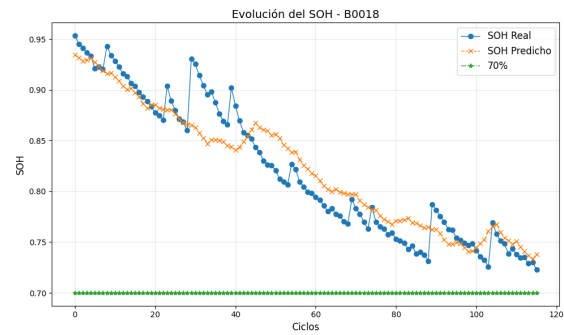
(a) Test battery B0005



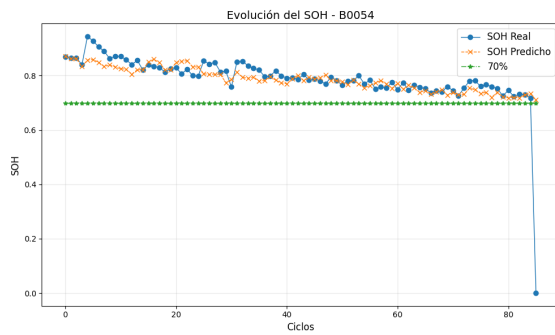
(b) Test battery B0006



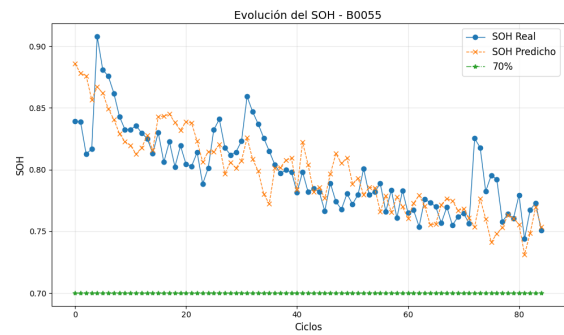
(c) Test battery B0007



(d) Test battery B0018



(e) Test battery B0054

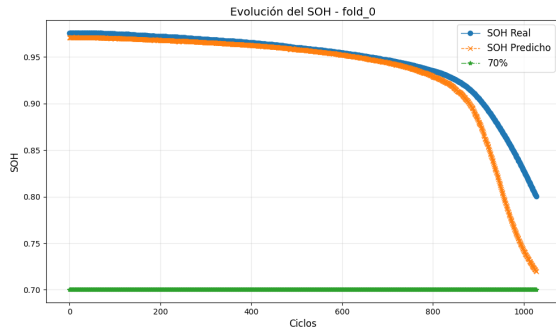


(f) Test battery B0055

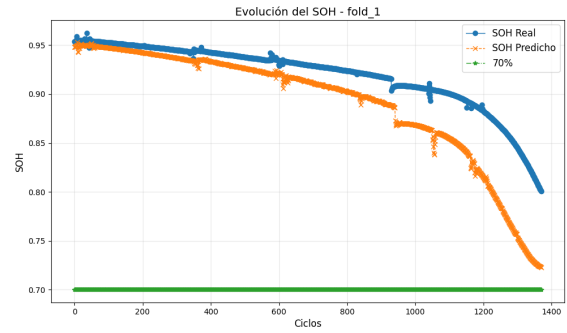
Figura A.3: Resultados de los modelos NASA

A.6. Resultados modelos MIT

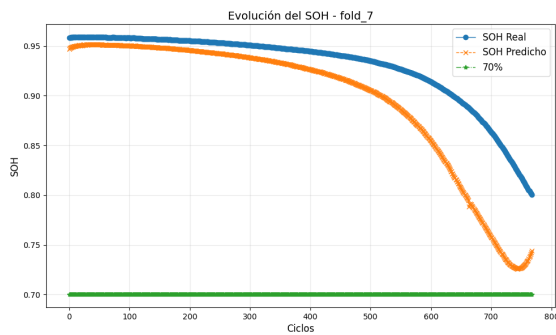
A.6.1. MSE como función de pérdida



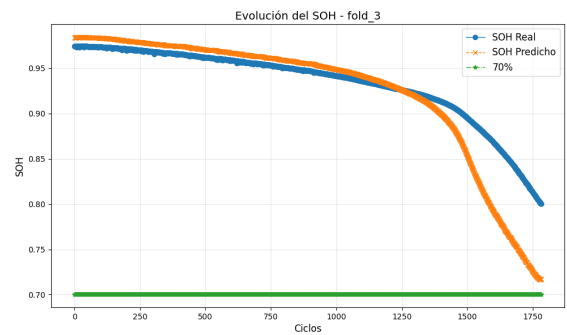
(a) Ejemplo - Fold 0



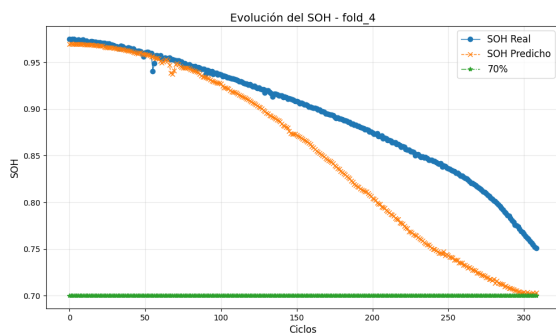
(b) Ejemplo - Fold 1



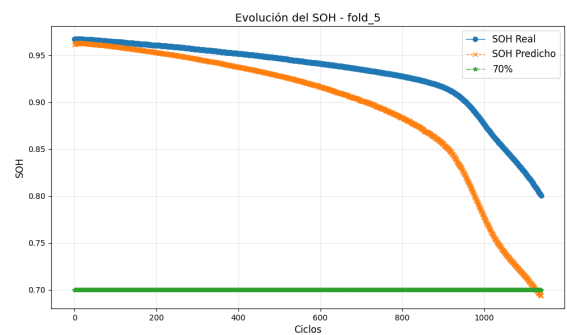
(c) Ejemplo - Fold 7



(d) Ejemplo - Fold 3



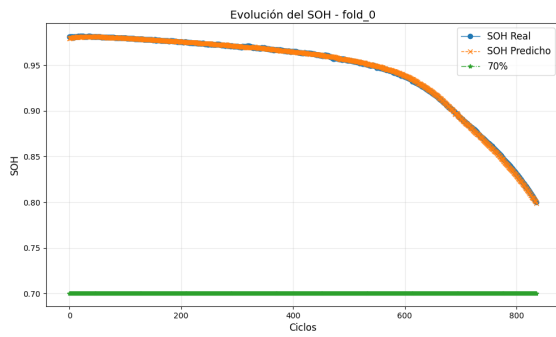
(e) Ejemplo - Fold 4



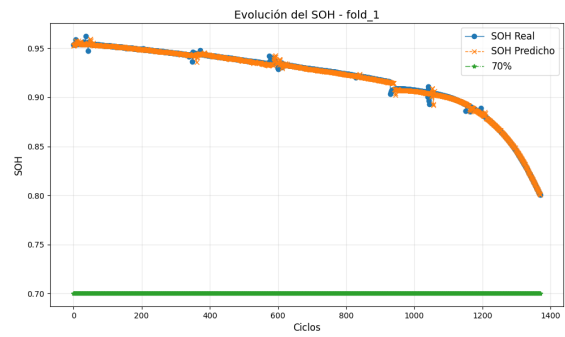
(f) Ejemplo - Fold 5

Figura A.4: Resultados de los modelos MIT con función de pérdida MSE

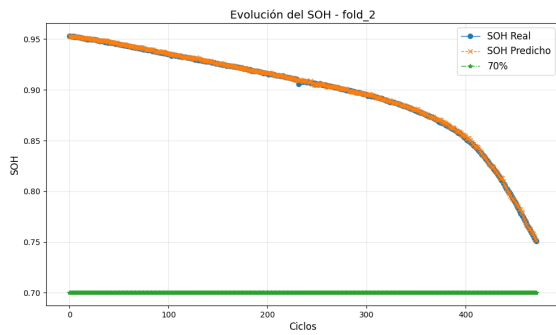
A.6.2. MSE ponderado como función de pérdida



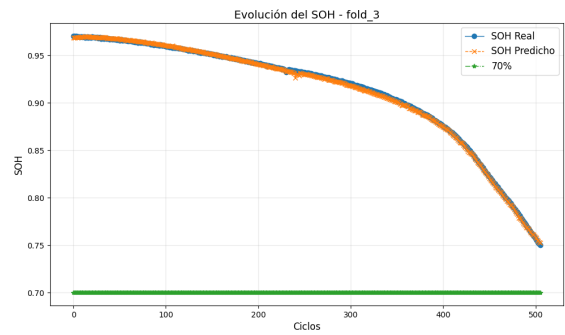
(a) Fold 0 - Test battery 25



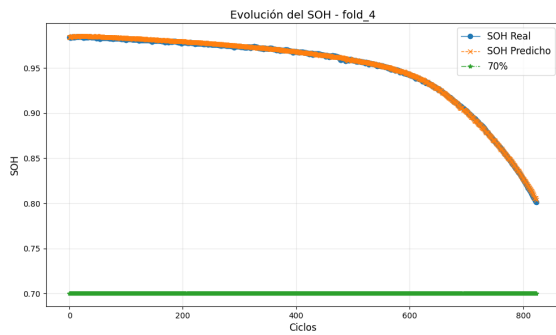
(b) Fold 1 - Test battery 83



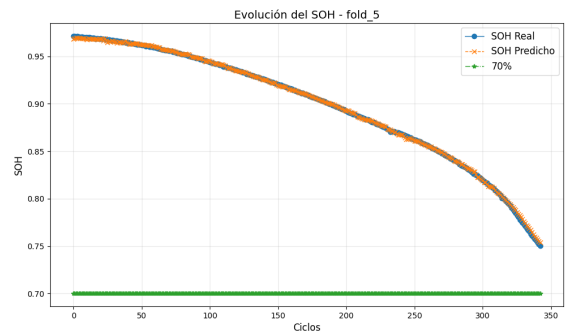
(c) Fold 2 - Test battery 111



(d) Fold 3 - Test battery 110

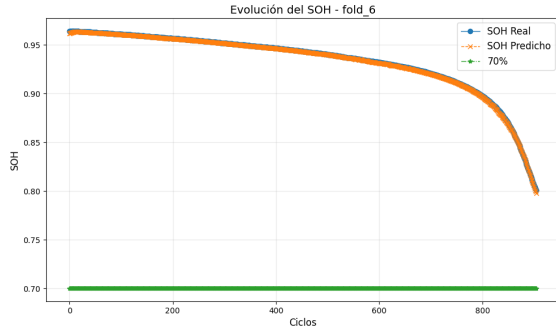


(e) Fold 4 - Test battery 27

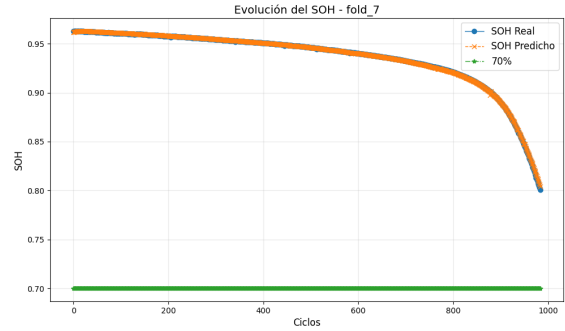


(f) Fold 5 - Test battery 95

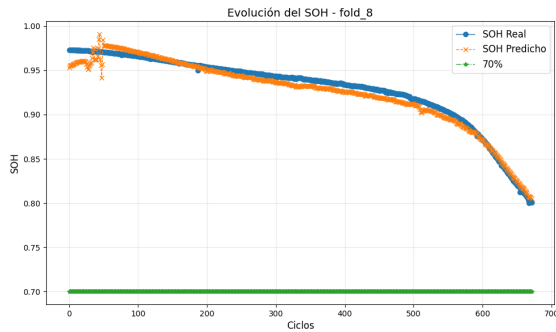
Figura A.5: Resultados de los modelos MIT (Parte 1)



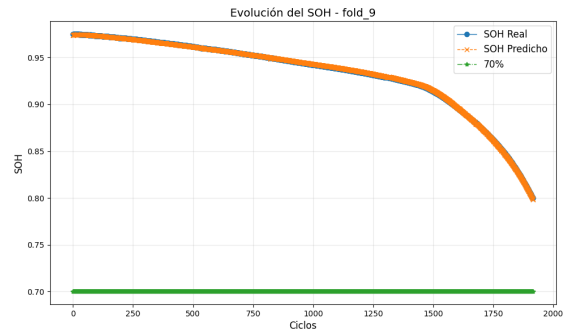
(g) Fold 6 - Test battery 82



(h) Fold 7 - Test battery 68



(i) Fold 8 - Test battery 18

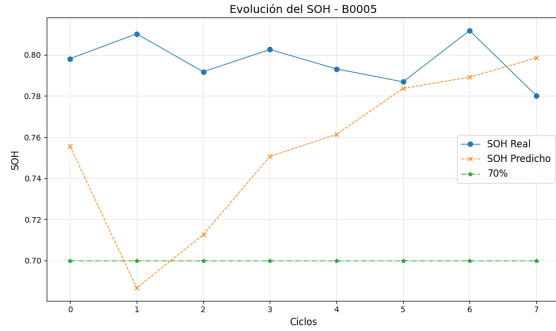


(j) Fold 9 - Test battery 84

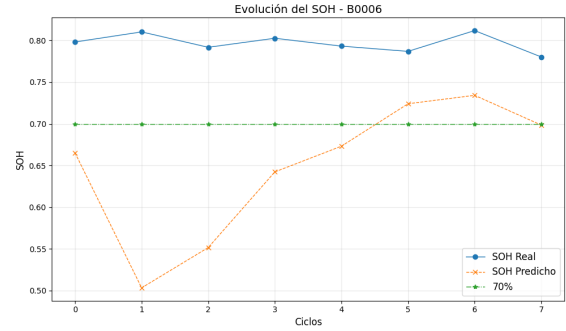
Figura A.5: Resultados de los modelos MIT (Parte 2)

A.7. Resultados modelo NASA con datos reales

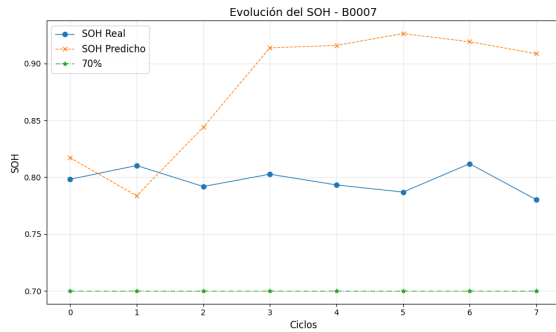
Resultados de los modelos NASA aplicados a datos reales del vehículo. Notar que el nombre del modelo corresponde a la batería que fue utilizada para validar en el esquema leave-one-out.



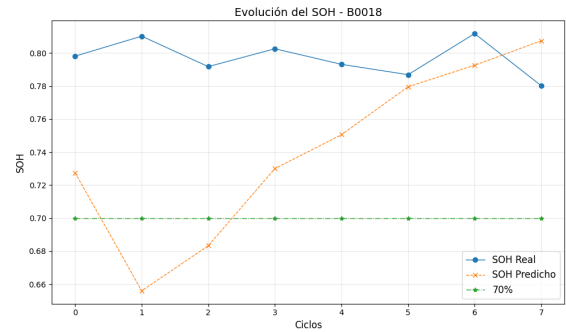
(a) Modelo B0005 en datos reales



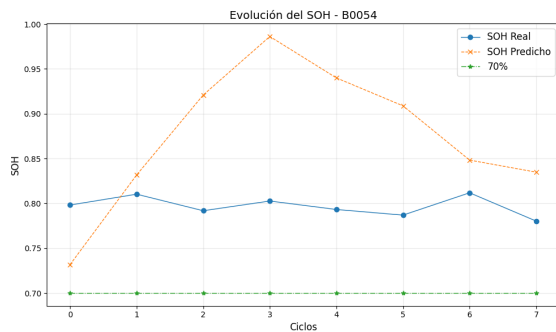
(b) Modelo B0006 en datos reales



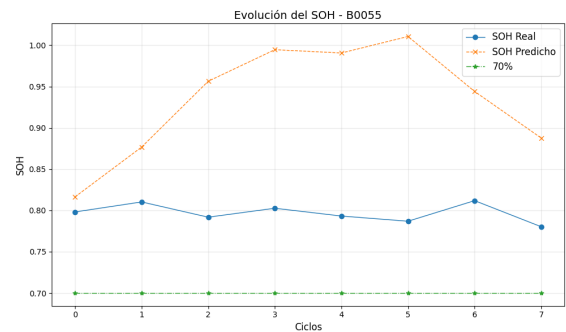
(c) Modelo B0007 en datos reales



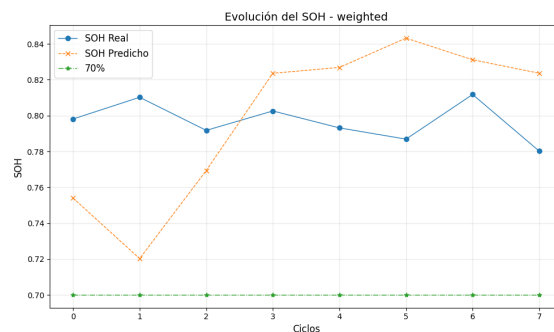
(d) Modelo B0018 en datos reales



(e) Modelo B0054 en datos reales



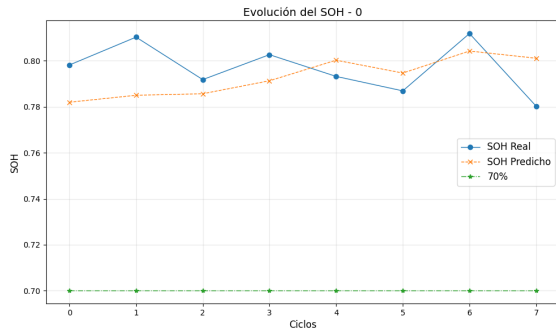
(f) Modelo B0055 en datos reales



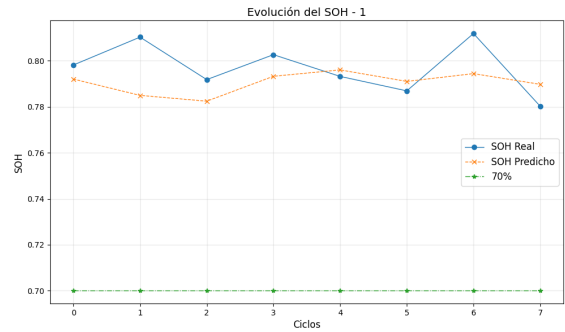
(g) Modelo ensamble ponderado en datos reales

Figura A.6: Resultados de los modelos NASA aplicados a datos reales del vehículo

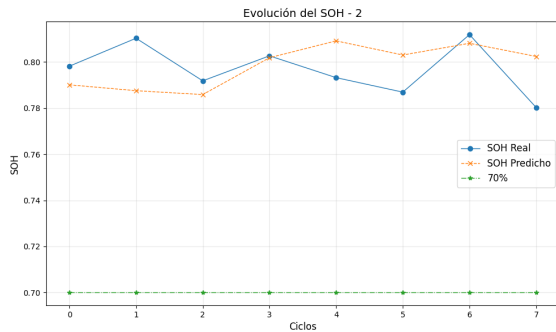
A.8. Resultados modelo MIT con datos reales



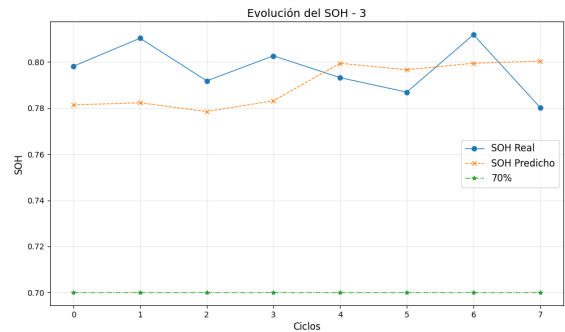
(a)



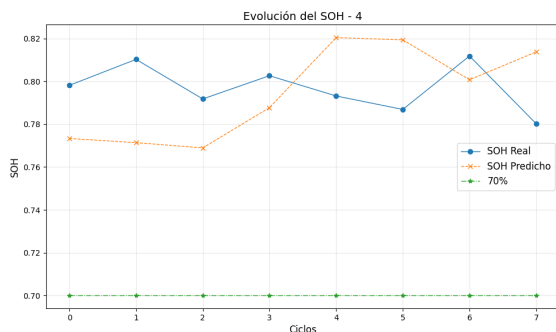
(b)



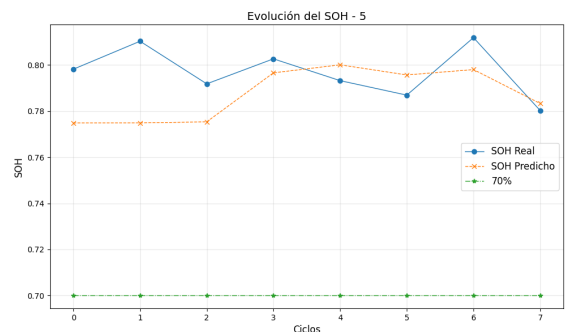
(c)



(d)

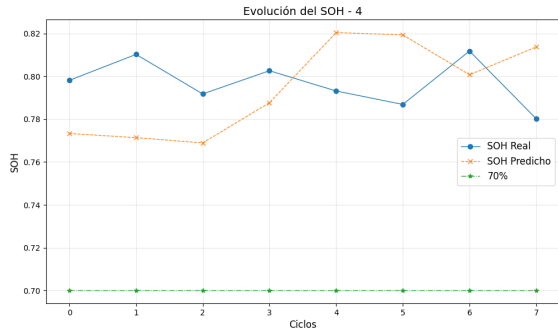


(e)

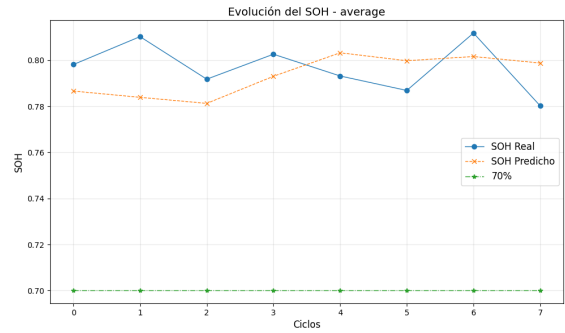


(f)

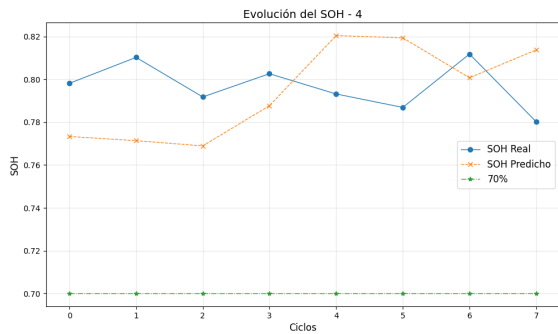
Figura A.7: Resultados de los modelos aplicados a datos reales del vehículo (Parte 1)



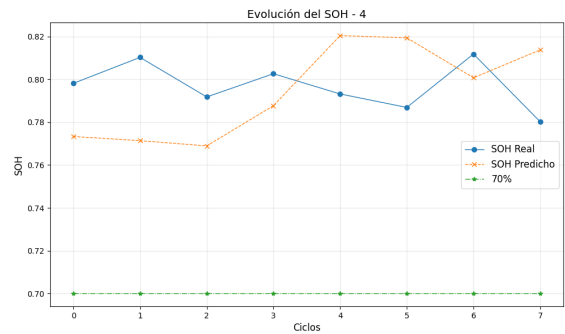
(g)



(h)

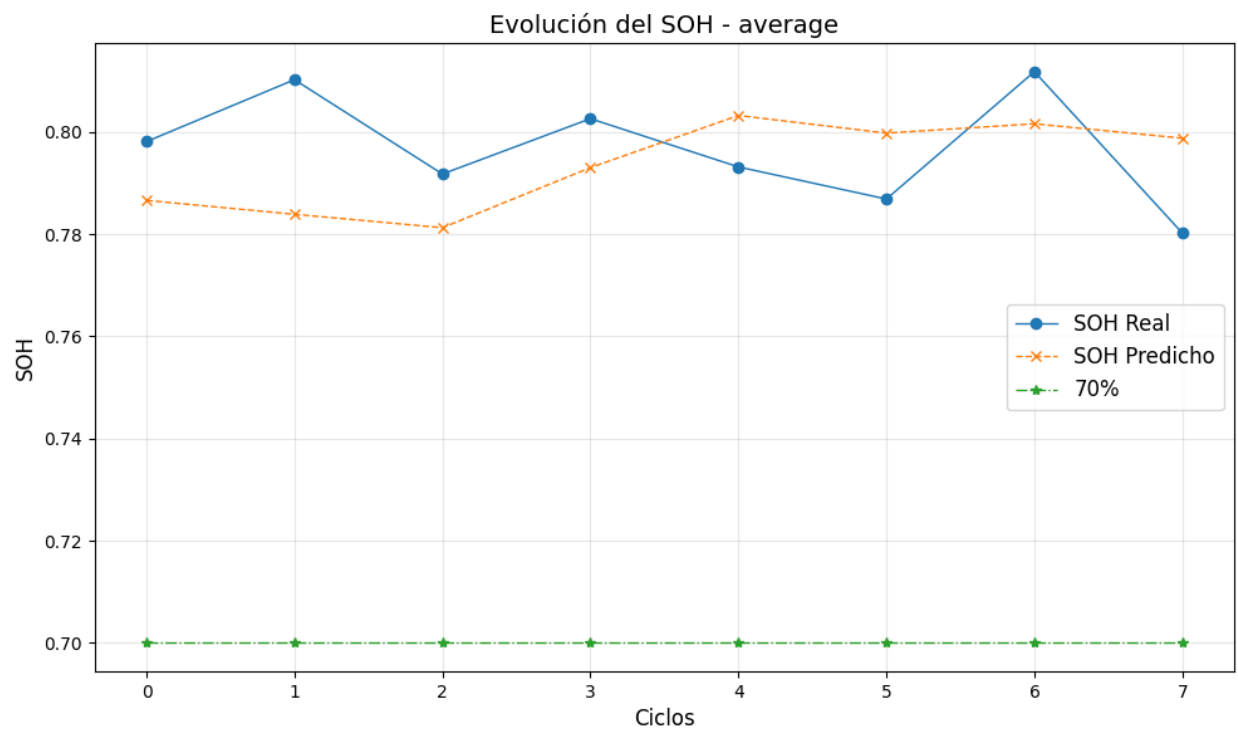


(i)



(j)

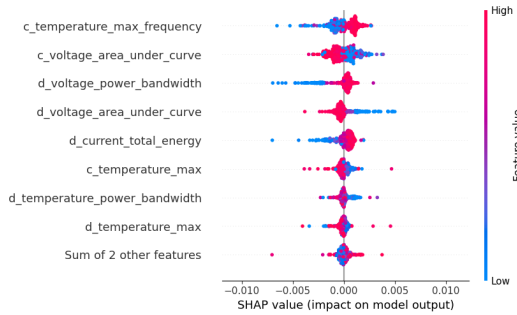
Figura A.7: Resultados de los modelos aplicados a datos reales del vehículo (Parte 2)



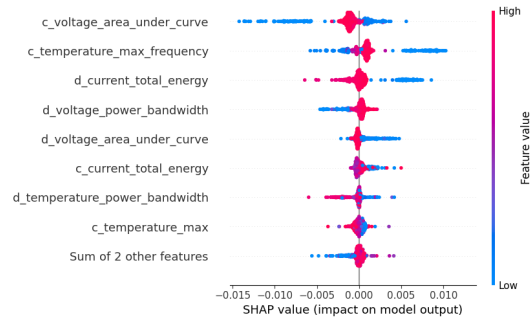
(k) Ensemble del promedio de las predicciones sobre datos reales

Figura A.7: Resultados de los modelos aplicados a datos reales del vehículo (Parte 3)

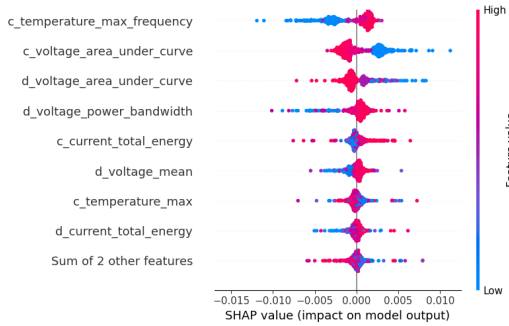
A.9. Gráficos *beeswarm* del análisis de SHAP



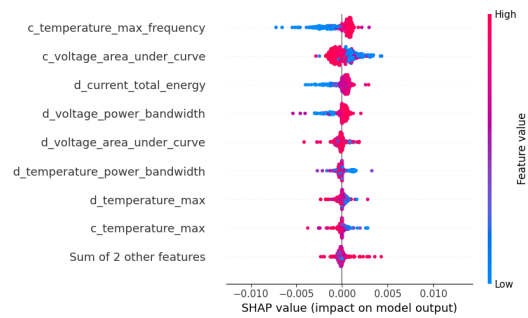
(a) *Beeswarm Fold 1*



(b) *Beeswarm Fold 2*

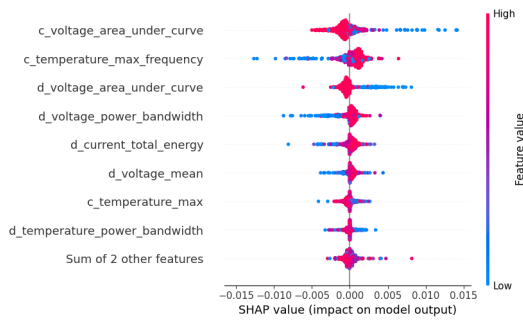


(c) *Beeswarm Fold 3*

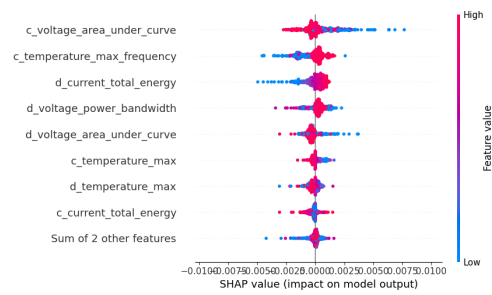


(d) *Beeswarm Fold 5*

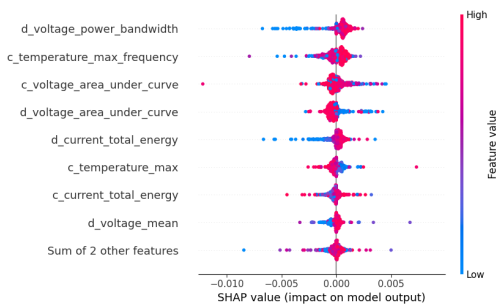
Figura A.8: Gráficos *beeswarm* del análisis de SHAP por *fold* (Parte 1)



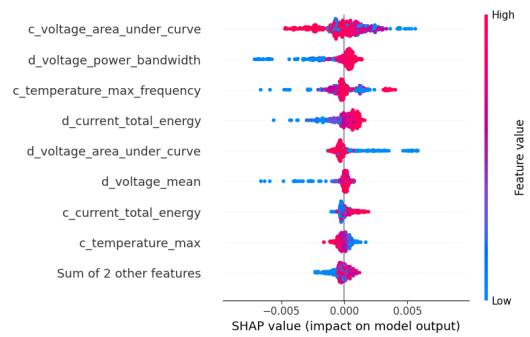
(e) *Beeswarm Fold 7*



(f) *Beeswarm Fold 8*



(g) *Beeswarm Fold 9*



(h) *Beeswarm Fold 10*

Figura A.8: Gráficos *beeswarm* del análisis de SHAP por *fold* (Parte 2)