



Máster en Big Data y Analítica Avanzada

Proyecto Final de Máster

**Modelado del Comportamiento de UFLS  
mediante Redes Neuronales Adaptadas a Datos  
Censurados**

Autor del proyecto:  
Alberto Fernández Carrillo

Director del proyecto:  
Mohammad Rajabdorri



Firma del autor



Firma del director

Madrid  
Junio 2025



## Resumen

Este trabajo aborda el modelado del comportamiento del esquema de **Desconexión Automática de Carga por Baja Frecuencia (UFLS)** mediante técnicas de aprendizaje automático adaptadas a **datos censurados**. El UFLS es una herramienta crítica en la operación de sistemas eléctricos, cuyo objetivo es preservar la estabilidad tras contingencias graves mediante el deslastre escalonado de carga. Sin embargo, su comportamiento presenta una alta no linealidad, una lógica condicional basada en umbrales y ecuaciones diferenciales que modelan la dinámica de los sistemas eléctricos. Por otro lado, también resulta difícil estimarlo mediante métodos comunes de aprendizaje automático basados en datos, ya que el UFLS derivado de distintos incidentes presenta una estructura censurada, al ser cero en muchos de ellos.

Para abordar estos retos, se ha desarrollado un modelo basado en redes neuronales, entrenado sobre datos que incluyen tanto eventos con actuación de UFLS como otros en los que no se produce deslastre (UFLS cero). El conjunto de datos presenta censura inferior estructural, ya que muchas observaciones con respuesta cero no implican ausencia de fenómeno, sino que se encuentran por debajo del umbral de activación. Se implementan funciones de pérdida especializadas (como el **MSE censurado**, el **MAE censurado** y formulaciones derivadas del **modelo Tobit**), que permiten al modelo adaptarse a esta estructura.

Los resultados muestran que el modelo predice con alta precisión la magnitud de la carga desconectada, especialmente al emplear funciones de pérdida que incorporan explícitamente la censura. Asimismo, se realiza una comparación con modelos lineales, que evidencian un rendimiento inferior. Finalmente, se discute la posibilidad de integrar el modelo en esquemas de optimización, y se señalan posibles líneas futuras de investigación.

## Agradecimientos

Quisiera expresar mi más sincero agradecimiento a todas las personas que, de una forma u otra, han contribuido al desarrollo de este trabajo y me han acompañado a lo largo de este año tan intenso y enriquecedor.

En primer lugar, a mi familia, por su apoyo incondicional, su paciencia y su confianza en mí durante todo este proceso. En especial, gracias por estar siempre ahí, incluso cuando no lo decía. Y, por supuesto, también a mi perro Bruce, que ha sabido darme compañía, tranquilidad y descanso en los momentos más tensos.

A mi tutor, Mohammad Rajabdorri, por su guía, su dedicación y sus valiosas orientaciones técnicas a lo largo del proyecto. Su experiencia ha sido clave para darle forma y profundidad a este trabajo.

A todos los compañeros y compañeras que he conocido este año durante el máster, por las conversaciones, el compañerismo y el ambiente que hemos construido. Vuestra presencia ha hecho que este año me enriquezca enormemente no solo en lo formativo, sino también en lo personal.

Y, por último, a mis amigos de siempre, por no dejar que desconectara del todo del mundo exterior, ni del mus, ni del Racing de Santander; y recordarme que hay vida más allá del Big Data.

A todos, gracias.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación del Proyecto . . . . .	1
1.2. Objetivos del Proyecto . . . . .	3
1.2.1. Objetivos Específicos . . . . .	3
<b>2. Estado del arte</b>	<b>5</b>
2.1. Estado del arte y aplicaciones del aprendizaje automático en UFLS	5
2.1.1. Fundamentos del esquema UFLS . . . . .	5
2.1.2. Estado de la cuestión: modelado del sistema eléctrico . . . . .	6
2.1.3. Aplicaciones de aprendizaje automático en UFLS . . . . .	7
2.2. Tratamiento de datos censurados y truncados en el modelado de UFLS . . . . .	8
2.2.1. Naturaleza de la censura y truncamiento en esquemas UFLS	8
2.2.2. Implicaciones estadísticas y sesgo en el aprendizaje automático	10
2.2.3. Enfoques estadísticos ante la censura . . . . .	11
2.2.4. Aplicabilidad en sistemas eléctricos . . . . .	12
<b>3. Aplicación y Validación de la Metodología</b>	<b>13</b>
3.1. Análisis de los datos . . . . .	13
3.2. Métodos y modelos desarrollados . . . . .	18
3.2.1. Arquitectura del modelo . . . . .	18
3.2.2. Funciones de pérdida utilizadas . . . . .	19
3.2.3. Entrenamiento . . . . .	21
3.3. Análisis de los resultados . . . . .	22
3.3.1. Evaluación cuantitativa . . . . .	22
3.3.2. Evaluación gráfica . . . . .	23
3.3.3. Evaluación complementaria . . . . .	27
<b>4. Conclusión</b>	<b>31</b>
4.1. Conclusiones generales del proyecto . . . . .	31



# Índice de figuras

3.1. Distribución ordenada de los valores de UFLS, destacando la censura inferior. . . . .	14
3.2. Distribuciones marginales de las variables del conjunto de datos. . .	15
3.3. Matriz de corr. de Pearson, entre las variables del conjunto de datos.	16
3.4. Representación conjunta de las variables mediante histogramas, dispersión y densidades. Casos con UFLS nula (negro) y con actuación (rojo). . . . .	17
3.5. Predicciones del modelo, según la función de pérdida. . . . .	23
3.6. Histogramas de los valores predichos, según la función de pérdida. .	24
3.7. Gráficos de residuos, según la función de pérdida. . . . .	26
3.8. Comparativa de la red neuronal frente a la regresión lineal, ambas con pérdida Tobit. . . . .	27



# Índice de cuadros

3.1. Estadísticos descriptivos de las variables del conjunto de datos. . . .	13
3.2. Rendimiento del modelo neuronal según la función de pérdida. . . .	22
3.3. Comparativa detallada entre modelos lineales y redes neuronales para cada métrica y función de pérdida. . . . .	28



# Capítulo 1

## Introducción

### 1.1. Motivación del Proyecto

La estabilidad de la frecuencia en los sistemas eléctricos es un aspecto crítico para garantizar la continuidad del suministro de energía y la seguridad de la infraestructura eléctrica; siendo cada vez más difícil de garantizar ese equilibrio en vista del cada vez mayor protagonismo de las fuentes de energía renovables en el mix energético y la menor inercia que esto conlleva [1, 2]. En este contexto, los sistemas de **Desconexión Automática de Carga por Baja Frecuencia (UFLS)**, por sus siglas en inglés) representan una herramienta fundamental para evitar colapsos generalizados ante contingencias graves. El funcionamiento de estos sistemas de protección se basa en la desconexión escalonada de carga cuando la frecuencia cae por debajo de ciertos umbrales predefinidos. A pesar de su importancia, su integración efectiva en modelos de planificación y operación del sistema eléctrico sigue siendo limitada, principalmente porque calcularlos en el proceso de programación operativa es complicado.

Los modelos UFLS tradicionales se rigen bajo reglas fijas y estructuras deterministas que reflejan la dinámica del sistema de manera discreta. Debido a estas características, es difícil representarlo en esquemas de optimización como el *unit commitment*, uno de los problemas clásicos de planificación eléctrica en el que se determina qué unidades generadoras deben estar activas en cada instante, considerando múltiples restricciones técnicas, costes y condiciones del sistema [3]. Incorporar de manera eficiente la lógica del UFLS en estos marcos permitiría no solo mejorar la calidad de las decisiones operativas, sino también utilizar estos mecanismos como una fuente adicional de reserva activa.

En este proyecto, se propone una solución alternativa basada en técnicas de aprendizaje automático, con el objetivo de desarrollar un modelo predictivo que

estime de forma eficiente la respuesta del esquema UFLS. A diferencia de los enfoques analíticos tradicionales, los modelos de aprendizaje automático permiten capturar relaciones no lineales complejas sin requerir una formulación explícita del sistema físico subyacente. Esto facilita su aplicación en entornos de optimización, donde la rapidez y la capacidad de generalización resultan fundamentales. También se han propuesto métodos para incorporar redes neuronales en problemas de programación lineal entera mixta (MILP, por sus siglas en inglés), lo que permite aplicarlos a problemas de optimización de programación, que habitualmente se resuelven como MILP [4].

Un aspecto distintivo de esta iniciativa radica en el tipo de datos disponibles para el entrenamiento de los modelos. Debido al diseño por umbrales del UFLS, gran parte de los eventos registrados no activan ninguna etapa del esquema, lo que se traduce en una elevada proporción de observaciones con respuesta nula. Estos datos se consideran **censurados**, ya que la variable de interés, la potencia desconectada, no refleja necesariamente su valor real, sino una ausencia de activación por no haberse alcanzado el umbral mínimo. Además, pueden existir restricciones físicas o técnicas que limiten el rango de valores observables, lo cual introduce truncamiento en los datos.

La presencia simultánea de censura y truncamiento plantea retos metodológicos importantes. Las funciones de pérdida tradicionales como el error absoluto medio (MAE) o el error cuadrático medio (MSE) no están diseñadas para este tipo de situaciones, y pueden conducir a estimaciones sesgadas o ineficientes [5]. Por ello, se hace necesario adoptar técnicas específicas que consideren la naturaleza incompleta de los datos. En particular, este proyecto explora funciones de pérdida derivadas de modelos estadísticos censurados, como el modelo Tobit, así como variantes reparametrizadas y adaptadas al aprendizaje profundo.

El desarrollo de este trabajo responde a una necesidad concreta en el ámbito de la ingeniería eléctrica: modelar de forma precisa la respuesta de esquemas UFLS en presencia de datos censurados. Además del componente aplicado, el proyecto plantea un reto metodológico relevante, al requerir técnicas de *Machine Learning* capaces de aprender a partir de información incompleta y estructuralmente sesgada. En este contexto, se propone y evalúa el uso de funciones de pérdida adaptadas, con el objetivo de construir modelos más representativos y compatibles con entornos de optimización energética.

## 1.2. Objetivos del Proyecto

El propósito fundamental de este Trabajo de Fin de Máster es diseñar e implementar un modelo de aprendizaje automático capaz de estimar el comportamiento de un esquema UFLS ante distintas contingencias del sistema eléctrico. A diferencia de las aproximaciones clásicas, este modelo se construirá considerando explícitamente la naturaleza censurada y truncada de los datos, utilizando las ya mencionadas funciones de pérdida adaptadas para este tipo de problemática.

### 1.2.1. Objetivos Específicos

- **Caracterización de los datos UFLS:** Analizar la estructura de los datos obtenidos mediante simulaciones de eventos UFLS, identificando la presencia y el impacto de la censura y el truncamiento.
- **Revisión de funciones de pérdida convencionales y adaptadas:** Estudiar las limitaciones del uso de funciones de pérdida estándar (MAE, MSE) en contextos con datos incompletos, y explorar funciones específicas como la pérdida Tobit o variantes censuradas del error absoluto y cuadrático.
- **Diseño del modelo predictivo:** Implementar un modelo de red neuronal profunda para la estimación de la potencia desconectada, integrando funciones de pérdida adaptadas a la censura y al truncamiento.
- **Adaptación y reimplementación de técnicas del estado del arte:** Basarse en los enfoques propuestos por Dănăilă et al.[5] para desarrollar y adaptar funciones de pérdida que permitan el entrenamiento eficaz de modelos en presencia de datos censurados.
- **Evaluación comparativa:** Realizar un análisis empírico del desempeño del modelo con distintas configuraciones de funciones de pérdida sobre un conjunto de datos simulados, evaluando la precisión, estabilidad y capacidad de generalización del modelo.



# Capítulo 2

## Estado del arte

### 2.1. Estado del arte y aplicaciones del aprendizaje automático en UFLS

#### 2.1.1. Fundamentos del esquema UFLS

El **Under-Frequency Load Shedding (UFLS)**, conocido en español como deslastre de carga por baja frecuencia, es un mecanismo de protección crítica empleado en sistemas eléctricos con el fin de mantener la estabilidad del sistema ante disturbios severos de frecuencia, evitando apagones generalizados [6]. El UFLS actúa como una medida correctiva, diseñada para desconectar automáticamente una cantidad de cargas (consumo eléctrico) previamente seleccionadas cuando la frecuencia del sistema cae por debajo de ciertos umbrales preestablecidos [7]. De este modo se busca reequilibrar rápidamente la generación y la demanda, impidiendo que la frecuencia continúe disminuyendo hasta niveles que puedan colapsar la red. La importancia del UFLS en la operación de redes modernas es tal que se considera una última línea de defensa para prevenir apagones totales (blackouts) durante contingencias como la pérdida repentina de generación eléctrica [8, 9].

En términos generales, la literatura sobre sistemas eléctricos identifica y clasifica los esquemas UFLS en tres categorías principales según su grado de adaptabilidad y complejidad operativa: convencionales, semi-adaptativos y adaptativos [10].

Los esquemas convencionales suelen estar divididos en etapas sucesivas, desconectando progresivamente diferentes bloques de carga conforme se alcanzan niveles decrecientes de frecuencia, asegurando así una respuesta escalonada ante eventos severos. Los esquemas convencionales, como su nombre indica, son los más utiliza-

dos actualmente debido a su simplicidad operativa, bajo coste y alta confiabilidad técnica. Además, las constantes variaciones en la tasa de cambio de la frecuencia (ROCOF) dificultan la implementación de los esquemas más avanzados que se describen más adelante [11]. Dentro de los UFLS convencionales, existe literatura sobre distintos esquemas posibles. Amraee et al. [12] utilizan un enfoque probabilístico que incorpora incertidumbres del sistema, mientras que Laghari et al. [13] proponen optimizar la cantidad de carga desconectada mediante un método basado en prioridades. Por su parte, Lobato et al. [14] desarrollan una técnica robusta para ajustar parámetros de esquemas UFLS convencionales usando datos generados sintéticamente. En general, estos esquemas tradicionales son los más utilizados actualmente en sistemas eléctricos a nivel mundial.

Por otro lado, los esquemas semi-adaptativos añaden cierta flexibilidad a los esquemas convencionales al considerar ajustes dinámicos limitados. Estos ajustes se basan generalmente en mediciones adicionales como ROCOF. Aunque la incorporación de ROCOF permite una respuesta algo más ajustada a las condiciones operativas reales del sistema eléctrico (tal y como se plantea en [15]), estos esquemas aún dependen considerablemente de configuraciones obtenidas mediante simulaciones previas, análisis estadísticos y experiencia operativa acumulada [16].

En contraste, los esquemas UFLS adaptativos representan la categoría más avanzada en términos de flexibilidad operativa y capacidad de respuesta dinámica. Estos sistemas ajustan automáticamente y en tiempo real la cantidad específica de carga que debe ser desconectada según las condiciones operativas actuales del sistema eléctrico [17, 18, 19, 20]. Para ello, requieren la implementación de sofisticados sistemas de medición y monitoreo continuo que proporcionen información detallada sobre el estado operativo del sistema, incluyendo variables como frecuencia, tasa de cambio de frecuencia, tensión y magnitud del desequilibrio generación-demanda. No obstante, la implementación práctica de estos esquemas adaptativos presenta limitaciones técnicas y económicas considerables, principalmente asociadas con los requisitos elevados de infraestructura de comunicaciones rápidas, confiables y seguras [11].

### 2.1.2. Estado de la cuestión: modelado del sistema eléctrico

Ya explicada la importancia del UFLS como mecanismo de protección, debe remarcar que su incorporación explícita en modelos de planificación operativa ha sido históricamente poco explorada en la literatura. La mayoría de las investigaciones previas han priorizado el desarrollo e implementación de medidas preventivas, tales como [21], que emplea un modelo *data-driven* con redes neuronales para pre-

decir el punto mínimo de frecuencia e incorporarlo en una formulación MILP. Por otro lado, en [22] se propone coordinar el UFLS con servicios auxiliares, utilizando mediciones en tiempo real para frenar la caída de frecuencia y evitar desconexiones innecesarias. No obstante, algunos estudios han comenzado a abordar esta integración de manera más directa y explícita.

Estos estudios recientes proponen modelos analíticos detallados que incluyen restricciones específicas para eventos UFLS, enfatizando la necesidad de considerar adecuadamente estas acciones correctivas en la planificación operativa. Esto se vuelve especialmente crítico en redes con una alta penetración de fuentes de energía renovables (solar o eólica) y una inercia reducida (como puede ser en sistemas insulares), donde las caídas de frecuencia se vuelven más abruptas e impredecibles. En este tipo de sistemas eléctricos los eventos UFLS resultan prácticamente inevitables ante perturbaciones importantes [23, 24, 9]. En este sentido, la literatura actual destaca que la consideración adecuada del UFLS puede permitir una reducción significativa en los requerimientos de reservas operativas, ofreciendo potenciales ahorros económicos y mejoras en la eficiencia global del sistema [25].

Cabe destacar que existe un creciente consenso en la literatura reciente sobre la importancia de usar métodos analíticos avanzados y *data driven* para la optimización de estos sistemas. El razonamiento detrás de este enfoque es que, mediante estos algoritmos *data-driven*, es posible anticipar y optimizar diversos aspectos de los sistemas eléctricos.

### 2.1.3. Aplicaciones de aprendizaje automático en UFLS

En los últimos años, han emergido técnicas basadas en aprendizaje automático (AA) aplicadas a los sistemas eléctricos, incluyendo el contexto del UFLS. Varios enfoques han sido explorados, especialmente mediante técnicas supervisadas para estimar directamente la magnitud de carga a desconectar. Ejemplo de ello, es el trabajo de Rajabdorri et al. [9], donde proponen generar datos sintéticos de UFLS ante distintas contingencias utilizando un modelo dinámico de frecuencia, y entrenar árboles de regresión capaces de predecir la cantidad de carga a desconectar en función de la potencia perdida, la inercia del sistema y las reservas de potencia disponibles. Además, el estudio incorpora un modelo **Tobit** [26] —analizado en mayor detalle más adelante— para tratar la **censura** inherente al problema, ya que el deslastre tiene un límite inferior de cero. El modelo aprendido se integra en un esquema de optimización de la operación (formulado como un problema MILP) para planificar reservas y mejorar la respuesta en frecuencia. Los resultados demuestran que el uso de AA permite estimar con precisión el UFLS requerido, reduciendo costes y mejorando la seguridad del sistema. Esta precisión puede incrementarse

aún más mediante técnicas más sofisticadas, como las redes neuronales profundas.

Otro enfoque relevante son los métodos de aprendizaje por refuerzo (RL) en la toma de decisiones secuenciales para UFLS. A diferencia de los enfoques tradicionales, estos algoritmos entrenan un agente que aprende a decidir cuánto y cuándo desconectar carga, buscando estabilizar la frecuencia con el menor sacrificio posible. Este aprendizaje se realiza mediante simulaciones repetidas de contingencias. Por ejemplo, [8] presenta un agente de RL que, tras un entrenamiento intensivo, logra estabilizar la frecuencia más rápido y con menos carga desconectada que los esquemas basados en umbrales. El agente es capaz de aprender políticas de actuación más eficientes, adaptándose mejor a cada situación del sistema.

Sin embargo, un desafío clave de los métodos *data-driven* es la necesidad de contar con datos de entrenamiento representativos de distintas contingencias y configuraciones del sistema eléctrico. En la práctica, los grandes disturbios de frecuencia (como pérdidas severas de generación) son poco frecuentes, por lo que se recurre habitualmente a simulaciones para generar escenarios de entrenamiento [27, 9]. Además, la capacidad de generalización es limitada: un modelo entrenado en un sistema concreto puede no funcionar igual de bien en otro con distinta topología o mix de generación, lo que ha motivado el uso de *transfer learning* y datos sintéticos variados. También se plantea el reto de la interpretabilidad, ya que modelos basados en *deep learning* o *reinforcement learning* suelen ser cajas negras difíciles de analizar. Aun así, los beneficios son destacables. Por ejemplo, [28] propone una estrategia que combina estimaciones no paramétricas de variabilidad eólica con un agente de RL entrenado en línea, logrando reducir más del 50 % la carga desconectada frente a esquemas UFLS tradicionales, manteniendo la frecuencia en márgenes seguros incluso bajo alta penetración renovable. Esto demuestra la mayor eficiencia y robustez que las técnicas de aprendizaje automático pueden aportar al UFLS moderno.

## 2.2. Tratamiento de datos censurados y truncados en el modelado de UFLS

### 2.2.1. Naturaleza de la censura y truncamiento en esquemas UFLS

En problemas del sistema eléctrico como el modelado de esquemas UFLS, los datos disponibles no siempre reflejan con fidelidad el comportamiento subyacente del sistema. En particular, al analizar la cantidad de carga desconectada tras una contingencia, es frecuente encontrarse con conjuntos de datos donde muchas

observaciones tienen valor cero. Sin embargo, este valor observado no siempre representa la verdadera magnitud del fenómeno, sino que puede ser consecuencia de la lógica de activación del propio esquema de protección. Esta situación es típica de problemas con datos **censurados**.

En términos generales, la **censura** ocurre cuando la variable de interés no puede observarse directamente por caer fuera de ciertos umbrales, aunque existe de forma latente. Formalmente, si denotamos por  $y^*$  la variable latente —el valor verdadero que habría tomado la respuesta sin restricciones— y por  $y$  la variable efectivamente observada, el modelo censurado se define como:

$$y = \begin{cases} v_l, & \text{si } y^* \leq \tau_l \\ y^*, & \text{si } \tau_l < y^* < \tau_u \\ v_u, & \text{si } y^* \geq \tau_u \end{cases} \quad (2.1)$$

Donde:

- $y^*$  es la variable latente o verdadera.
- $y$  es la variable observada.
- $\tau_l$  y  $\tau_u$  son los umbrales inferior y superior de censura, respectivamente.
- $v_l$  y  $v_u$  son los valores sustituidos cuando  $y^*$  no es observable (por debajo de  $\tau_l$  o por encima de  $\tau_u$ ).

Este esquema permite representar tanto censura inferior como superior, y es flexible en el sentido de que los valores sustituidos ( $v_l$ ,  $v_u$ ) no tienen por qué coincidir con los umbrales de censura ( $\tau_l$ ,  $\tau_u$ ). En particular, el caso más común de censura inferior en problemas energéticos se presenta cuando los valores reales por debajo de cierto umbral se registran como cero, lo cual puede introducir un sesgo significativo en el análisis si no se trata adecuadamente [5].

La censura no es un fenómeno exclusivo del UFLS, pues existen múltiples contextos donde está presente. A continuación se proponen dos ejemplos de censura, para ilustrar los dos tipos que existen. Un caso típico es el de los límites de detección en mediciones en laboratorios, donde los instrumentos de medida —al igual que en el UFLS— tienen un umbral mínimo (límite inferior de cuantificación), cualquier valor real por debajo de ese límite no se mide exactamente sino que se reporta como “Por debajo del límite”. Así, podríamos saber que la concentración de un contaminante es inferior, digamos, a 0.5 ppm, pero ignoramos su concentración real exacta (podría ser 0.3, 0.1, etc.) . Ese dato está **censurado por la izquierda**, puesto que solo tenemos la información de que el valor verdadero es menor

que cierto umbral. De forma análoga, existe la **censura por la derecha** que suele darse en ámbitos asociados a privacidad y recopilación de datos, sobre todo para bases de datos socioeconómicas: con el fin de proteger la privacidad, los ingresos muy altos se informan solo hasta un tope (p.e.: "por encima de 100,000 euros") en vez del valor exacto, lo cual implica que los valores reales por encima de ese límite quedan censurados (se desconoce cuánto exactamente ganan las personas cuyo ingreso supera el umbral declarado) [5].

Un fenómeno relacionado pero conceptualmente distinto es el *truncamiento*. A diferencia de la censura, en el truncamiento los valores que caen fuera de un intervalo dado no solo no se observan, sino que ni siquiera se incluyen en el conjunto de datos. Es decir, la observación completa se descarta. Por ejemplo, si el sistema de medición o simulación impone límites estructurales (como no registrar eventos con carga desconectada menor a un cierto umbral de criticidad), esos datos se pierden por truncamiento. Este fenómeno afecta la composición misma del dataset y requiere un tratamiento estadístico distinto. Aunque ambos fenómenos pueden coexistir, este trabajo se enfoca exclusivamente en la censura, ya que los datos utilizados no presentan evidencia de truncamiento sistemático [5].

En el caso específico del modelado de esquemas UFLS abordado en este trabajo, los datos presentan **censura inferior**. La lógica del sistema establece que no se activa ningún escalón del UFLS si la frecuencia no cae por debajo del primer umbral técnico de actuación, lo cual suele corresponder a una desconexión de al menos 1 MW. No obstante, los valores registrados para estos casos son  $v_l = 0$ , lo que implica que cualquier valor latente  $y^*$  menor que  $\tau_l \approx 1$  será observado como cero. Esta condición genera una acumulación artificial de observaciones nulas que, si no se trata adecuadamente, puede distorsionar tanto la estimación como la interpretación del modelo predictivo. La consideración explícita de la censura en el entrenamiento del modelo resulta, por tanto, esencial para obtener predicciones coherentes con el comportamiento real del sistema.

### 2.2.2. Implicaciones estadísticas y sesgo en el aprendizaje automático

La presencia de datos censurados supone un desafío significativo al entrenar modelos de aprendizaje automático y estadístico. El problema central es que no se dispone del valor real completo de ciertas muestras, lo que rompe los supuestos tradicionales de muchos algoritmos. En un escenario estándar de aprendizaje supervisado (por ejemplo, regresión), asumimos que para cada instancia de entrenamiento tenemos un valor objetivo claro para predecir. Con datos censurados, en cambio, algunas instancias no tienen un valor objetivo concreto sino un intervalo

(en nuestro caso., “el valor es menor que 1”). Si se ignorase la censura y se tratase ese dato como si valiera exactamente el valor asociado al intervalo censurado (volviendo a nuestro problema, asignar el valor 0 ante el caso anterior), se introduciría un sesgo sistemático: estaríamos subestimando consistentemente los verdaderos valores que eran mayores. Por otro lado, eliminar todas las observaciones censuradas del conjunto de datos tampoco es una buena solución, ya que se pierde información valiosa y se puede sesgar la muestra (especialmente si la censura no ocurre al azar y su presencia aporta información).

En términos de modelado, la censura implica que la función de verosimilitud o los criterios de pérdida deben modificarse para incorporar la información parcial disponible. Un algoritmo de aprendizaje debe aprovechar el hecho de saber, por ejemplo, que  $y > y_0$  aun sin conocer el valor exacto de  $y$ . Esto complica tanto el entrenamiento como la evaluación de los modelos, pues no se puede calcular un error cuadrático medio (MSE) convencional entre una predicción  $\hat{y}$  y un valor  $y$  verdadero cuando  $y$  no es un número concreto sino, digamos, “mayor que  $y_0$ ”. En resumen, los métodos tradicionales de ajuste directo minimizando el error observación a observación no se aplican de forma directa. Se requieren técnicas especializadas para evitar estimaciones sesgadas y para extraer toda la información posible de los datos censurados. Manejar adecuadamente la censura es crucial para lograr modelos confiables y precisos; de hecho, en muchas aplicaciones es preferible adoptar métodos que consideren explícitamente la censura en lugar de intentar “rellenar” o ignorar estos datos, ya que un tratamiento inadecuado puede llevar a conclusiones erróneas [5].

### 2.2.3. Enfoques estadísticos ante la censura

En el campo del análisis de datos censurados existen dos líneas metodológicas destacadas cuya combinación resulta especialmente relevante para el enfoque planteado en este trabajo. Por un lado, los enfoques estadísticos clásicos —como el modelo **Tobit** [26]— ofrecen una base sólida gracias a su formulación explícita de la censura mediante máxima verosimilitud, permitiendo utilizar toda la información disponible sin introducir sesgos significativos. Este tipo de modelos y sus extensiones paramétricas y no paramétricas han sido ampliamente utilizados en econometría y análisis de supervivencia, y proporcionan un marco interpretativo riguroso. Por otro lado, los desarrollos recientes han resaltado la utilidad de técnicas de redes neuronales profundas adaptadas a datos censurados. En concreto, Dănăilă y Buiu [5] proponen dos funciones de pérdida especializadas —la **CMAE** (Censored Mean Absolute Error) y la **CMSE** (Censored Mean Squared Error)— diseñadas para entrenar redes en presencia de censura inferior, donde los valores por debajo del umbral se registran como ceros. Estas funciones permiten que el

modelo penalice los errores únicamente en la parte no censurada de los datos, logrando combinar robustez y precisión sin ignorar los ceros censurados. En su estudio, observaron que, si bien la pérdida basada en Tobit ofrecía un rendimiento ligeramente superior, las variantes CMAE y CMSE resultaban más simples de implementar y también muy efectivas. Estos conceptos se analizarán más en detalle en la sección 3.2.

#### 2.2.4. Aplicabilidad en sistemas eléctricos

La capacidad de modelar correctamente datos censurados y truncados no es solo un reto técnico, sino un habilitador clave para la integración del UFLS en herramientas de planificación y operación. Un modelo entrenado con funciones adaptadas puede representar con mayor fidelidad el comportamiento real del sistema, y además ser transformado en una estructura compatible con programación matemática (por ejemplo, como una red neuronal representable por restricciones tipo MILP).

Esto abre la puerta a esquemas híbridos donde la lógica aprendida se incorpora directamente en procesos de optimización, permitiendo decisiones más informadas y coherentes con la física del sistema. En consecuencia, el presente proyecto adopta estas metodologías como base para el diseño de su modelo predictivo, con el objetivo de superar las limitaciones de los enfoques tradicionales y avanzar hacia una integración operativa eficiente del UFLS [9].

# Capítulo 3

## Aplicación y Validación de la Metodología

### 3.1. Análisis de los datos

Antes de entrenar el modelo predictivo, se realizó un análisis exploratorio del conjunto de datos con el objetivo de identificar distribuciones, relaciones entre variables y posibles sesgos estructurales. En la Tabla 3.1 se presenta un resumen estadístico de las variables más relevantes: inercia total del sistema, ganancia ponderada del modelo turbina-gobernador ( $K$ ), potencia perdida ante contingencias, reserva disponible y respuesta observada del esquema UFLS.

Estadístico	Inercia (MW·s)	K pond. (MW)	Potencia perd. (MW)	Reserva (MW)	UFLS (MW)
<b>count</b>	3990	3990	3990	3990	3990
<b>mean</b>	92.55	929.57	5.10	10.12	2.26
<b>std dev</b>	15.57	143.11	1.91	2.93	2.29
<b>min</b>	39.26	450.00	2.50	0.50	0.00
<b>25 %</b>	87.74	876.00	3.50	8.50	0.00
<b>50 %</b>	93.28	932.00	4.50	10.50	1.74
<b>75 %</b>	101.87	1013.00	7.00	12.00	4.82
<b>max</b>	133.18	1327.00	10.00	18.00	6.94

Cuadro 3.1: Estadísticos descriptivos de las variables del conjunto de datos.

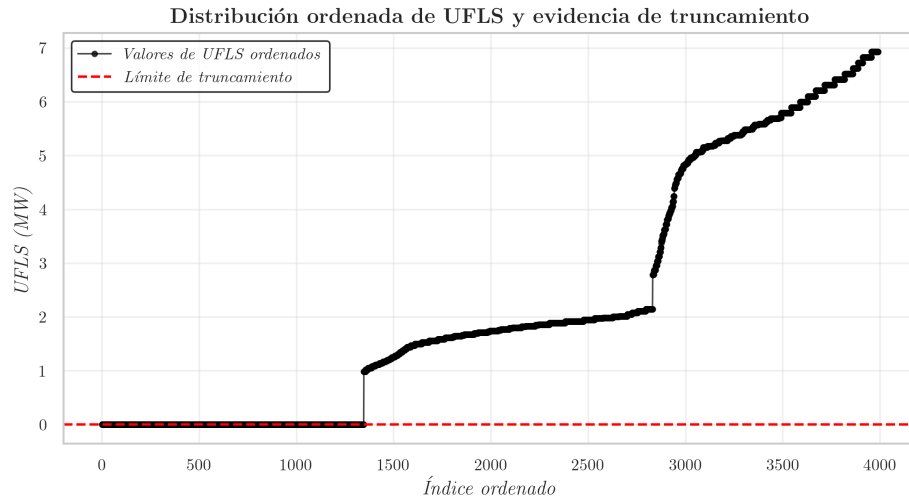


Figura 3.1: Distribución ordenada de los valores de UFLS, destacando la censura inferior.

Como se aprecia en la Figura 3.1, una fracción considerable del conjunto de datos presenta valores de UFLS exactamente iguales a cero, lo que evidencia la presencia de censura inferior. Estos ceros no necesariamente representan situaciones en las que no se ha producido deslastre de carga, sino que reflejan casos en los que la frecuencia del sistema no ha descendido por debajo del umbral de activación del esquema UFLS. Por tanto, el valor real de desconexión podría ser mayor que cero, pero permanece oculto por el diseño del sistema de protección. Esta característica introduce un sesgo importante en la distribución de los datos, que debe ser abordado explícitamente mediante técnicas estadísticas o de aprendizaje automático adaptadas al tratamiento de datos censurados.

Para estudiar la distribución marginal de cada variable se generaron histogramas, que permiten visualizar su forma y simetría. En particular, el histograma de la variable objetivo UFLS (Figura 3.2) vuelve a evidenciar el fenómeno de la censura inferior: se observa una gran acumulación de observaciones exactamente en cero, correspondientes a escenarios en los que no se activó el deslastre de carga. Por tanto, estos ceros deben interpretarse como observaciones censuradas, donde el valor real subyacente es desconocido pero pudiendo ser mayor que cero.

Además, se confirma que no existen valores negativos de UFLS en el conjunto de datos, lo cual es coherente con la naturaleza física del problema: no tiene sentido registrar valores negativos de potencia deslastrada. Esta característica justifica el uso de funciones de activación o restricciones que impidan predicciones por debajo

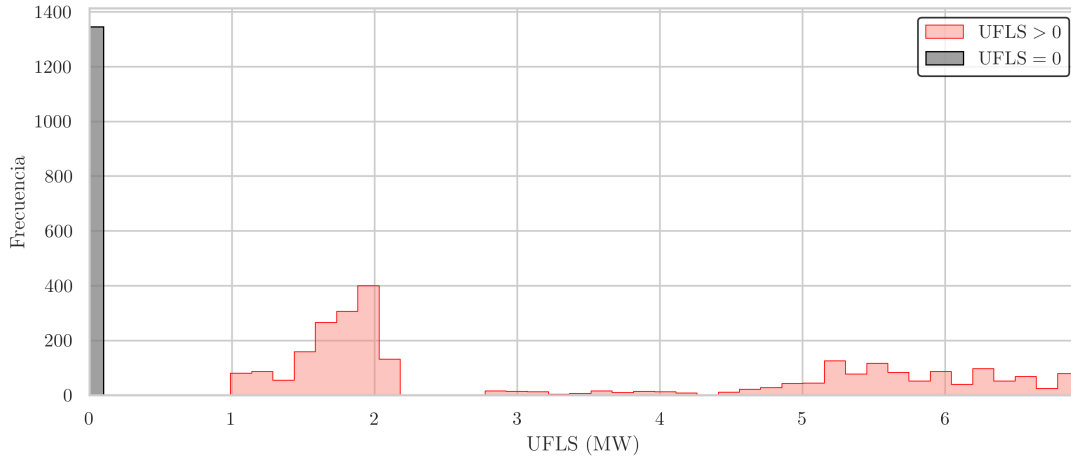


Figura 3.2: Distribuciones marginales de las variables del conjunto de datos.

de cero en los modelos entrenados, asegurando consistencia entre las salidas del modelo y los principios físicos del sistema eléctrico.

En la Figura 3.3 se muestra la matriz de correlación de Pearson entre las variables del conjunto de datos. Esta matriz revela relaciones esperables desde el punto de vista físico. Por ejemplo, la variable objetivo UFLS se correlaciona negativamente con variables como la inercia y la reserva, lo que refleja su efecto estabilizador sobre la frecuencia del sistema: cuanto mayor es la inercia o la capacidad de reserva, menor es la probabilidad de que se active el esquema de desconexión automática. Por el contrario, la correlación positiva con la potencia perdida (0.89) es coherente, ya que eventos con mayor pérdida de potencia tienden a activar con más frecuencia la protección UFLS.

Un aspecto notable es la altísima correlación entre las variables ‘inercia’ y ‘K ponderada’ (0.99), lo que sugiere una posible redundancia informativa. Tal y como se observa también en los estadísticos de la Tabla 3.1, estas variables presentan no solo un comportamiento muy similar, sino también una escala proporcional –siendo ‘K ponderada’ aproximadamente un orden de magnitud mayor–, lo que podría hacer pensar que son esencialmente la misma variable expresada en unidades distintas. Bajo este supuesto, se evaluó eliminar una de ellas para evitar colinealidad y simplificar el modelo. Sin embargo, las pruebas empíricas mostraron que al eliminar cualquiera de las dos, el rendimiento del modelo se deterioraba de forma consistente, tanto en métricas globales como en zonas censuradas. Esto sugiere que, a pesar de su correlación, ambas variables aportan información complementaria relevante en el contexto no lineal del modelo, posiblemente al in-

teractuar de forma distinta con el resto de entradas. Por tanto, se ha decidido mantener ambas variables en el conjunto de entrada.

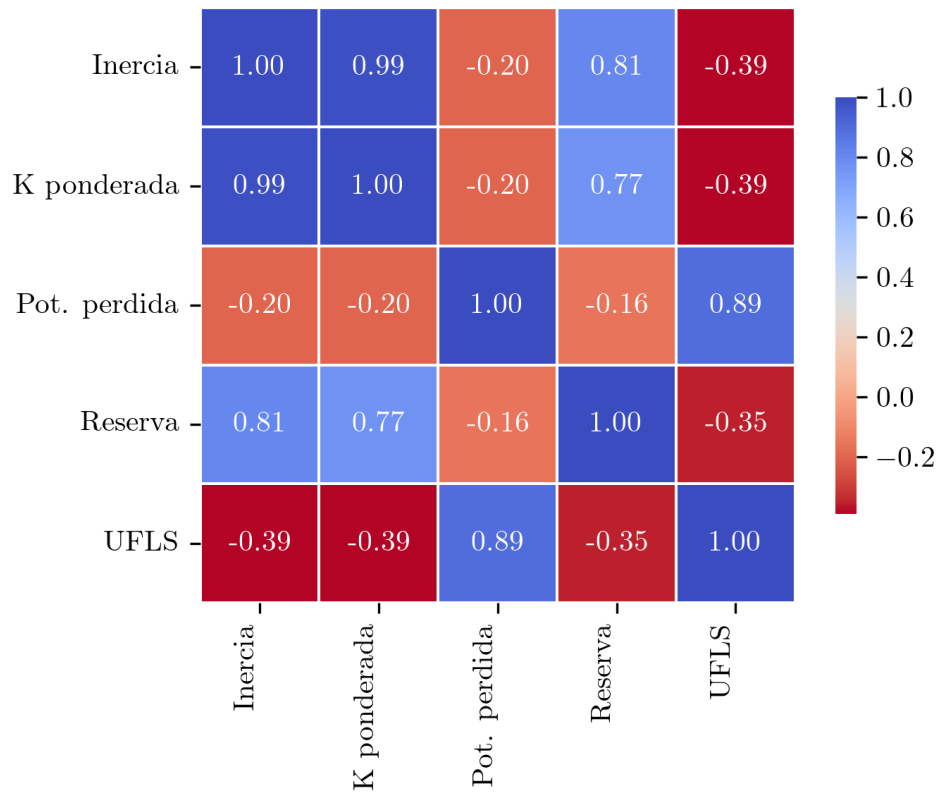


Figura 3.3: Matriz de corr. de Pearson, entre las variables del conjunto de datos.

La Figura 3.4 muestra un *pairplot* entre todas las variables del conjunto de datos, diferenciando con claridad los casos sin activación de UFLS (en negro) y aquellos con desconexión activa (en rojo). Se observan patrones de densidad y agrupamiento que revelan una fuerte capacidad discriminativa en algunas variables. En particular, la **potencia perdida** y la **reserva** destacan como predictores relevantes: los eventos con respuesta UFLS distinta de cero tienden a asociarse con valores altos de potencia perdida y niveles bajos de reserva, lo cual refleja correctamente la lógica del sistema eléctrico, donde un mayor desequilibrio entre generación y demanda combinado con poca capacidad de respuesta acelera la caída de frecuencia. Las curvas de densidad muestran separaciones nítidas entre ambos grupos, especialmente en la potencia perdida, donde los eventos con UFLS nulo están fuertemente concentrados en la parte baja del dominio.

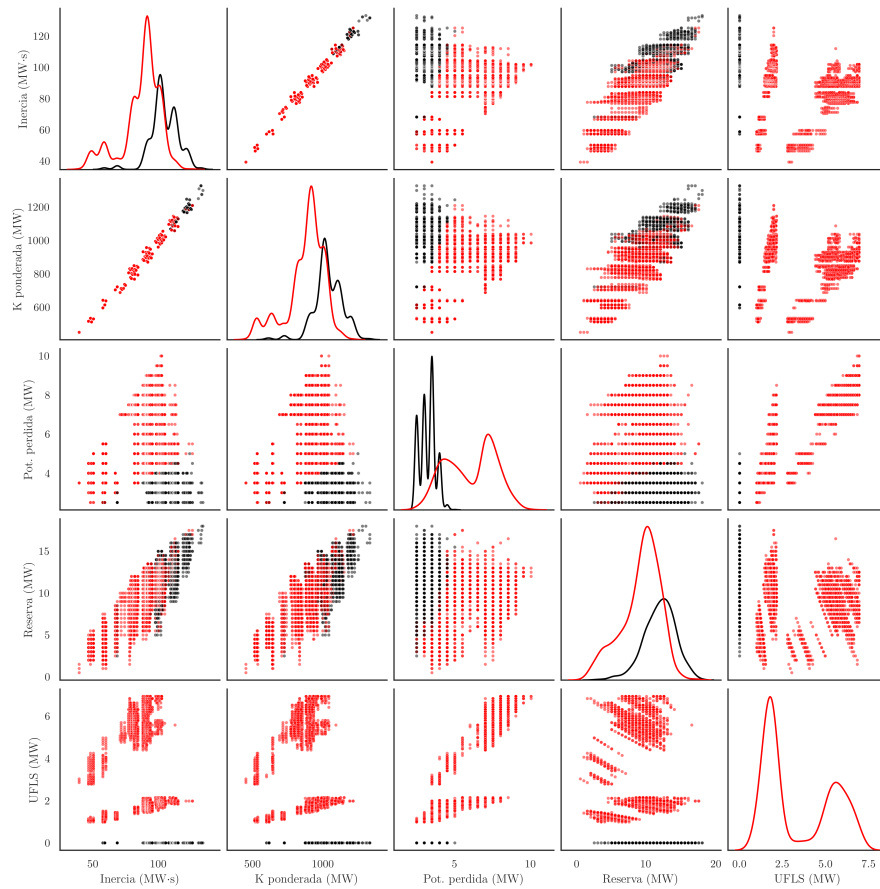


Figura 3.4: Representación conjunta de las variables mediante histogramas, dispersión y densidades. Casos con UFLS nula (negro) y con actuación (rojo).

Por otro lado, la **inercia** y la **K ponderada** muestran una correlación casi perfecta, ya anticipada en la matriz de correlación y atribuible a su relación lineal directa (K ponderada es esencialmente inercia escalada). No obstante, su capacidad para separar casos censurados y no censurados es más limitada. Finalmente, el panel de la variable objetivo UFLS vuelve a poner de manifiesto la presencia de **censura inferior**: una gran acumulación de observaciones en cero contrasta con una distribución más continua para el resto de valores, lo que confirma que no todas las observaciones nulas reflejan ausencia real de desconexión, sino también situaciones donde el valor subyacente no se ha podido observar debido a umbrales técnicos. Esta visualización respalda la necesidad de un tratamiento adecuado de la censura durante el modelado.

Este análisis preliminar de la información proporciona una base sólida para la

comprensión del comportamiento del conjunto de datos y justifica tanto la selección de variables como la posterior elección del modelo y su entrenamiento.

## 3.2. Métodos y modelos desarrollados

Como ya se ha mencionado previamente, el presente trabajo propone un enfoque basado en aprendizaje automático para modelar la respuesta del sistema UFLS, considerando explícitamente la naturaleza censurada de los datos. Para ello, se desarrolla una arquitectura neuronal que no solo estima el valor esperado de la respuesta, sino también su incertidumbre, integrando este conocimiento en funciones de pérdida adaptadas al contexto estadístico.

### 3.2.1. Arquitectura del modelo

El modelo principal empleado es `MLPSigmaV1`, una red neuronal totalmente conectada con una sección compartida que proyecta las variables de entrada a un espacio latente, desde el cual se generan dos salidas: `mu`, que representa la predicción del valor esperado, y `sigma`, una estimación de la desviación típica asociada. Ambas salidas se procesan con `Softplus` para garantizar valores positivos (pues como es lógico, un valor negativo de UFLS carece de sentido):

```
class MLPSigmaV1(nn.Module):
    def __init__(self, input_dim):
        super().__init__()
        self.shared = nn.Sequential(
            nn.Linear(input_dim, 32),
            nn.ReLU(),
            nn.Linear(32, 16),
            nn.ReLU()
        )
        self.mu_head = nn.Sequential(
            nn.Linear(16, 1),
            nn.Softplus()
        )
        self.sigma_head = nn.Sequential(
            nn.Linear(16, 1),
            nn.Softplus()
        )
```

Esta arquitectura permite abordar funciones de pérdida probabilísticas heterocedásticas como las derivadas del modelo Tobit, expuestas más adelante. Además,

se ha definido una función de penalización adicional que castiga las predicciones cercanas a cero en casos donde el valor real es claramente mayor, reduciendo el impacto de los llamados falsos ceros.

Esta penalización se activa cuando el valor predicho cae por debajo de un cierto umbral, pero el valor real es superior, y su efecto se modula mediante un peso configurable. Aunque podría parecer redundante respecto a las ya introducidas funciones de pérdida diseñadas para datos censurados –como CMSE, CMAE o la pérdida Tobit– en realidad actúa como un complemento específico, no como un sustituto. Como se explica más adelante, estas funciones modelan globalmente la censura mediante ajustes en el cálculo del error o mediante formulaciones probabilísticas, pero no penalizan de forma explícita los falsos ceros. Al introducir esta penalización adicional, se refuerza el aprendizaje en la frontera de activación del sistema UFLS sin interferir con las propiedades estadísticas del modelo base, contribuyendo a reducir errores críticos en la predicción de la carga deslastrada.

### 3.2.2. Funciones de pérdida utilizadas

Para abordar la naturaleza censurada de los datos del UFLS, se han evaluado varias funciones de pérdida, tanto convencionales como adaptadas. Las pérdidas convencionales (MAE y MSE) no tienen en cuenta la censura explícita, pero son útiles como línea base. Las funciones adaptadas, en cambio, introducen formulaciones que reflejan de forma explícita las regiones censuradas o truncadas del espacio de predicción.

Entre las pérdidas específicas se encuentran el **Censored Mean Absolute Error (CMAE)** y el **Censored Mean Squared Error (CMSE)**, que modifican las versiones tradicionales incorporando términos adicionales para cubrir observaciones truncadas. Estas formulaciones se definen como:

$$\text{CMAE}_T = \text{CMAE} + \sum_i |f_i(x_i) - t_l| \quad (3.1)$$

$$\text{CMSE}_T = \text{CMSE} + \sum_i (f_i(x_i) - t_l)^2 \quad (3.2)$$

Donde  $f_i(x_i)$  es la predicción del modelo para el ejemplo  $i$ , y  $t_l$  representa el umbral inferior de truncamiento. Estas funciones penalizan adicionalmente las predicciones en zonas donde no hay información completa, ayudando a mejorar la robustez del entrenamiento. La justificación teórica de su uso se apoya en trabajos recientes ya citados previamente sobre el modelado de variables censuradas mediante aprendizaje profundo supervisado [5].

Una alternativa probabilística ampliamente utilizada para datos censurados es la pérdida basada en el modelo **Tobit** [26]. En su versión no escalada (“unscalled”), la función de pérdida se define como la suma de tres componentes: el error cuadrático medio sobre los datos observados, la log-verosimilitud acumulada para las observaciones censuradas por la izquierda, y otra para las censuradas por la derecha. Esta formulación considera también la posible existencia de truncamiento, restando la contribución del área truncada:

$$\mathcal{L}_{\text{Tobit}} = \mathcal{L}_p + \mathcal{L}'_l + \mathcal{L}_u \quad (3.3)$$

Donde cada término se descompone como:

$$\mathcal{L}_p = \sum_i \left[ \frac{1}{2} \left( \frac{y_i - f(x_i)}{\sigma} \right)^2 + \ln(|\sigma| + \varepsilon) \right] \quad (3.4)$$

$$\mathcal{L}'_l = - \sum_j \ln \left( \Phi \left( \frac{c_l - f(x_j)}{\sigma} \right) - \Phi \left( \frac{t_l - f(x_j)}{\sigma} \right) + \varepsilon \right) \quad (3.5)$$

$$\mathcal{L}_u = - \sum_k \ln \left( \Phi \left( \frac{f(x_k) - c_u}{\sigma} \right) + \varepsilon \right) \quad (3.6)$$

Aquí  $\Phi$  es la función de distribución acumulada de una normal estándar,  $t_l$  y  $c_l$  son los umbrales de truncamiento y censura inferior, respectivamente, y  $c_u$  el umbral de censura superior.

Para tener en consideración la **heterocedasticidad**<sup>1</sup>, se utiliza una extensión en la que cada observación cuenta con su propia desviación típica  $\sigma_i$ . Esta versión ajusta tanto los términos de log-verosimilitud como las normalizaciones por  $\sigma_i$ :

$$\mathcal{L}_{\text{Tobit-Hetero}} = \mathcal{L}_p^{\text{hetero}} + \mathcal{L}'_l^{\text{hetero}} + \mathcal{L}_u^{\text{hetero}} \quad (3.7)$$

con

$$\mathcal{L}_p^{\text{hetero}} = - \sum_i \left( -\frac{1}{2} \left( \frac{y_i - \mu_i}{\sigma_i} \right)^2 - \log(\sigma_i + \varepsilon) \right) \quad (3.8)$$

$$\mathcal{L}'_l^{\text{hetero}} = - \sum_j \log \left( \Phi \left( \frac{y_j - \mu_j}{\sigma_j} \right) - \Phi \left( \frac{t_l - \mu_j}{\sigma_j} \right) + \varepsilon \right) \quad (3.9)$$

<sup>1</sup>La **heterocedasticidad** se refiere a la variación no constante del error en un modelo de regresión. Específicamente, ocurre cuando la varianza de los errores o residuos no es constante a lo largo de todas las observaciones del modelo.

$$\mathcal{L}_u^{\text{hetero}} = - \sum_k \log \left( \Phi \left( \frac{\mu_k - y_k}{\sigma_k} \right) - \Phi \left( \frac{\mu_k - t_u}{\sigma_k} \right) + \varepsilon \right) \quad (3.10)$$

Estas funciones se combinan con estrategias de entrenamiento personalizadas para garantizar una correcta propagación del gradiente incluso en regiones donde no se observa respuesta del sistema (cuando  $UFLS = 0$ ), permitiendo al modelo adaptarse tanto a la dinámica del sistema como a la lógica de censura estructural.

### 3.2.3. Entrenamiento

El entrenamiento se realiza con el optimizador Adam, con tasa de aprendizaje fija y 80 épocas. Durante cada iteración se computa la función de pérdida total, incluyendo –cuando procede– términos de log-verosimilitud para datos censurados y la penalización descrita anteriormente. La estructura del entrenamiento contempla la distinción entre funciones de pérdida heterocedásticas, parametrizadas y convencionales, adaptando la forma en que se pasan las salidas del modelo al cálculo de la pérdida.

Tras el entrenamiento, se evalúa el modelo utilizando varias métricas tradicionales de regresión:

- **MAE (Mean Absolute Error):** mide el promedio de los errores absolutos entre predicción y valor real, proporcionando una métrica sencilla y robusta frente a valores atípicos.
- **MSE (Mean Squared Error):** calcula el promedio del cuadrado de los errores, penalizando más los errores grandes y favoreciendo ajustes suaves.
- **NRMSE (Normalized Root Mean Squared Error):** normaliza el error cuadrático medio respecto al rango o media de la variable objetivo, permitiendo comparaciones entre modelos o datasets distintos.
- **$R^2$  (Coeficiente de determinación):** indica la proporción de la varianza de la variable dependiente explicada por el modelo; valores cercanos a 1 indican un buen ajuste.

Esta evaluación se realiza para cada una de las funciones de pérdida señaladas, permitiendo así comparar el rendimiento del modelo bajo diferentes tratamientos de la censura, demostrando la importancia de utilizar funciones de pérdida adaptadas para una predicción precisa y coherente con la estructura de los datos.

### 3.3. Análisis de los resultados

Para analizar el rendimiento del modelo entrenado con distintas funciones de pérdida, se presentan a continuación los resultados tanto cuantitativos como gráficos.

#### 3.3.1. Evaluación cuantitativa

En la Tabla 3.2 se resumen las métricas obtenidas por cada función de pérdida, incluyendo MAE, MSE, NRMSE,  $R^2$ .

Pérdida	MAE	MSE	NRMSE	$R^2$
CMAE	0.3662	0.5458	0.3202	0.8929
CMSE	0.4141	0.5063	0.3084	0.9071
Tobit	0.3957	0.4689	0.2968	0.9088
Tobit Hetero	0.5446	0.6555	0.3510	0.8714

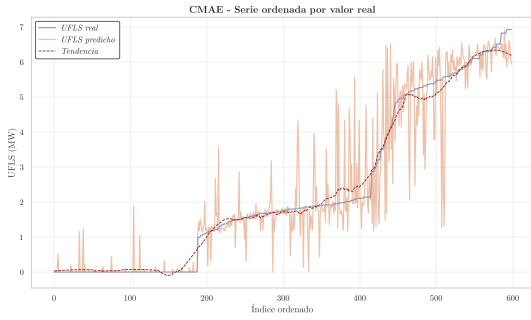
Cuadro 3.2: Rendimiento del modelo neuronal según la función de pérdida.

Los resultados expuestos en la Tabla 3.2 indican que la red propuesta obtiene su mejor rendimiento al ser entrenada con la pérdida **Tobit clásica**, alcanzando el menor error cuadrático medio ( $MSE = 0.4689$ ) y el mejor coeficiente de determinación ( $R^2 = 0.9088$ ), lo que evidencia una excelente capacidad para explicar la varianza de los datos.

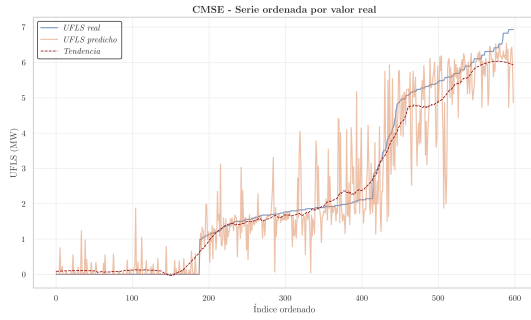
La pérdida **CMSE** presenta un comportamiento similar, con métricas también competitivas, lo que la convierte en una alternativa sólida y más sencilla de implementar. Aunque la pérdida **CMAE** logra el menor error absoluto medio ( $MAE = 0.3662$ ), su rendimiento global es inferior en términos de varianza explicada y error relativo, lo que sugiere que prioriza la robustez frente a outliers a costa de una menor precisión general.

Finalmente, la **versión heterocedástica del modelo Tobit** no mejora los resultados y, de hecho, muestra un deterioro claro en todas las métricas, probablemente debido a una mala estimación de la varianza o a un sobreajuste del modelo.

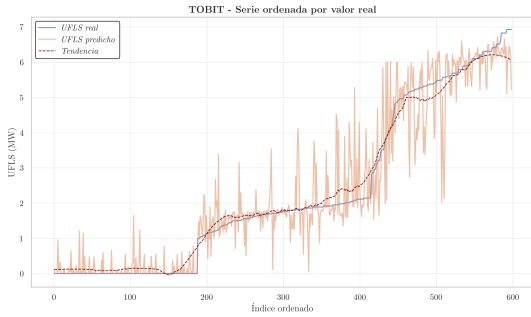
En conjunto, los resultados respaldan el uso de la **pérdida Tobit estándar como la opción más eficaz para esta arquitectura**.



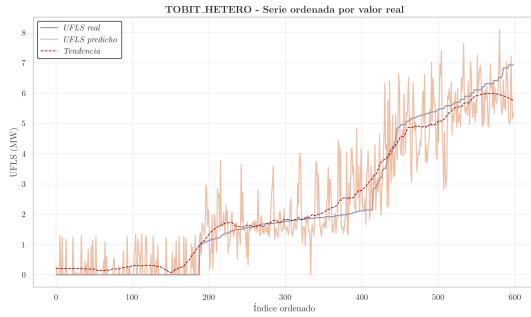
(a) Función de pérdida: CMAE



(b) Función de pérdida: CMSE



(c) Función de pérdida: Tobit



(d) Func. de pérdida: Tobit Heterocedást.

Figura 3.5: Predicciones del modelo, según la función de pérdida.

### 3.3.2. Evaluación gráfica

En las Figuras 3.5 a 3.7 se presentan los resultados gráficos en formato mosaico para cada función de pérdida. Se muestran, por ejemplo, gráficos de dispersión predicción vs real, histogramas de errores, y distribuciones de residuos.

Las gráficas presentadas en la Figura 3.5 muestran la calidad de las **predicciones** del modelo bajo distintas funciones de pérdida, permitiendo visualizar cómo se comporta la red a lo largo del rango de valores reales de UFLS. En todas las figuras, se representa la serie real con censura (línea azul), la predicción del modelo (línea naranja) y una tendencia suavizada (línea roja discontinua) que permite evaluar de forma clara el ajuste global.

El modelo entrenado con la pérdida **Tobit clásica** ofrece el comportamiento más equilibrado. La tendencia predicha se ajusta notablemente bien a la evolución de los valores reales, especialmente en las zonas donde hay transiciones suaves y aumento progresivo de UFLS. Aunque las predicciones crudas presentan cierta dispersión, el ajuste general es suave y coherente, lo que respalda su buen desempeño

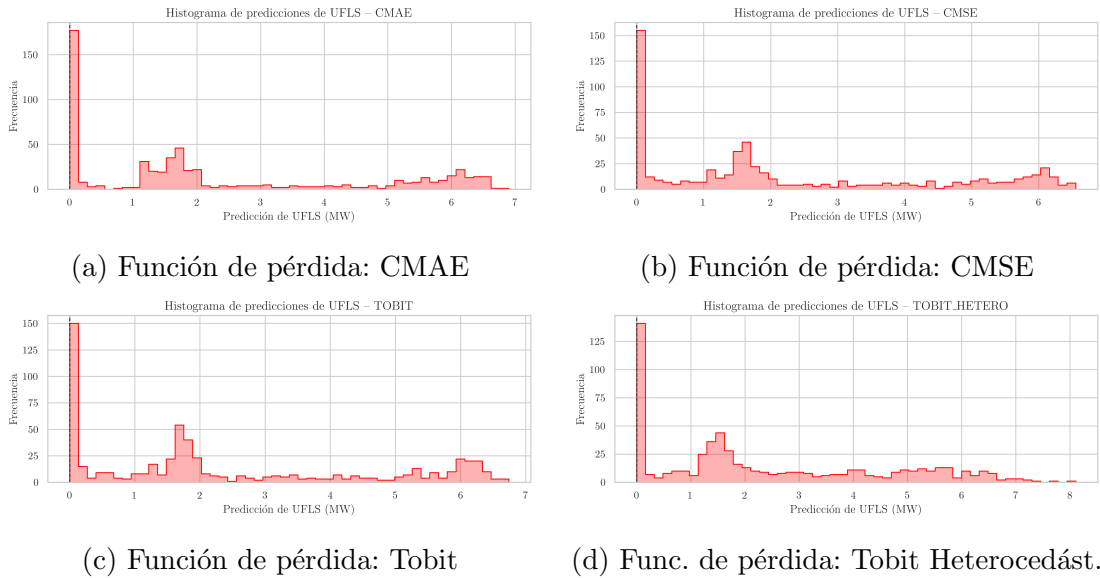


Figura 3.6: Histogramas de los valores predichos, según la función de pérdida.

observado también en las métricas cuantitativas. Esta opción logra capturar tanto el nivel como la forma de la señal sin introducir ruido adicional.

La pérdida **CMSE** también proporciona resultados competitivos, con una tendencia suavizada que sigue razonablemente bien la serie real, aunque con una ligera subestimación en los valores altos. Por su parte, la pérdida **CMAE** logra una buena aproximación general, pero presenta más variabilidad en zonas intermedias, reflejando su preferencia por errores absolutos pequeños. Finalmente, el modelo entrenado con la pérdida **Tobit heterocedástica** muestra mayor ruido y oscilaciones en la tendencia, posiblemente debido, de nuevo, a una estimación inestable de la varianza.

En conjunto, las gráficas de predicciones refuerzan la conclusión de que la pérdida Tobit clásica es la más adecuada para este problema, proporcionando predicciones coherentes, precisas y estables. La CMSE se presenta como una alternativa razonable, mientras que CMAE y Tobit hetero muestran limitaciones claras en términos de ajuste y robustez.

Los **histogramas** de predicciones correspondientes a las distintas funciones de pérdida, representados en la Figura 3.6, permiten analizar con mayor detalle cómo se distribuyen los valores estimados de UFLS, particularmente en relación con la censura desde cero que afecta a los datos. Una de las observaciones más relevantes

es la capacidad del modelo, especialmente cuando es entrenado con funciones como Tobit o CMSE, para generar predicciones distintas de cero en el rango comprendido entre 0 y 1. Este comportamiento es clave, ya que indica que la red ha aprendido a discernir cuándo un valor observado como cero no representa una ausencia real de deslastre, sino un dato censurado cuyo valor verdadero es simplemente inferior al umbral de detección.

Este efecto se traduce en una presencia significativa de valores predichos en la zona cercana al cero, lo cual sugiere que el modelo no se limita a replicar ciegamente la censura, sino que intenta estimar la magnitud real oculta. No obstante, también se observa una concentración notable de valores exactamente en cero, lo cual no es necesariamente malo. Esto se debe a que los datos reales pueden contener instancias sin ningún deslastre real de carga ( $UFLS = 0$  de forma legítima). En estos casos es razonable, a la par que deseable, que el modelo reproduzca también estos casos. En ese sentido, una proporción considerable de ceros en la predicción puede interpretarse como un reflejo correcto de la estructura estadística de los datos.

Otro aspecto positivo es la total ausencia de valores negativos en las predicciones. Este resultado valida el uso de restricciones específicas durante el entrenamiento, como funciones de activación tipo Softplus o penalizaciones personalizadas, que impiden que el modelo genere estimaciones físicamente inconsistentes. Recordemos que un valor negativo de UFLS carece de sentido práctico, ya que no es posible deslastrear una cantidad negativa de potencia. Por tanto, la red no solo ajusta bien la magnitud de los valores predichos, sino que también respeta las restricciones lógicas del problema, manteniendo todas sus salidas en el dominio válido.

En conjunto, los histogramas refuerzan la idea de que el modelo no solo es estadísticamente competente, sino que también es consciente de la naturaleza censurada del problema y capaz de dar una respuesta coherente, tanto numérica como interpretativamente, a los distintos rangos de valores de UFLS.

Los gráficos de **residuos** para cada función de pérdida permiten observar el comportamiento de los errores en función de la predicción de UFLS. En general, CMSE y Tobit vuelven a ser las funciones que demuestran mejor rendimiento, presentando una mayor concentración de residuos cerca de cero, lo que sugiere un mejor ajuste.

El modelo con pérdida Tobit destaca por presentar un patrón muy centrado en torno al eje horizontal, sin grandes asimetrías ni bandas de error sistemáticas. Los residuos son pequeños y distribuidos uniformemente, lo que indica una estimación

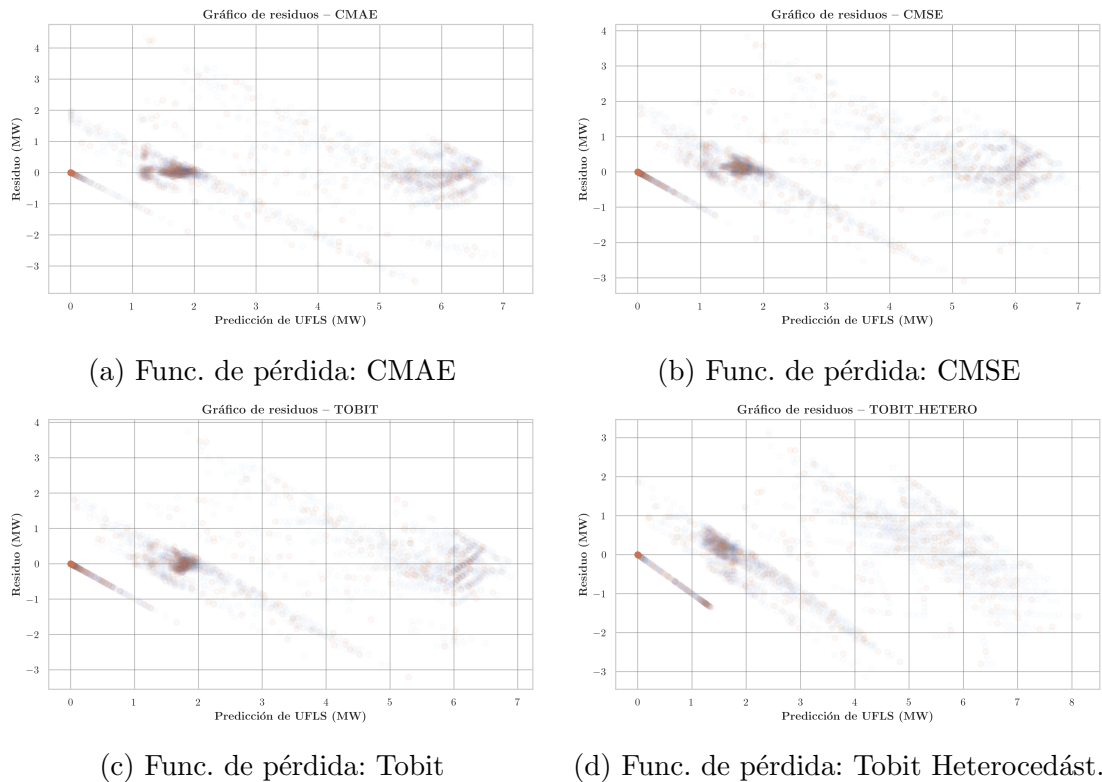


Figura 3.7: Gráficos de residuos, según la función de pérdida.

eficaz y coherente del comportamiento del sistema.

El modelo CMSE también muestra un patrón bien estructurado, con una mayor dispersión que el Tobit pero sin errores extremos destacables. En cambio, el modelo CMAE presenta una nube de residuos más irregular y una mayor concentración en algunos tramos, lo que podría estar asociado a su sensibilidad reducida frente a errores individuales grandes. Finalmente, el modelo Tobit heterocedástico genera los residuos más dispersos y menos centrados, lo cual es coherente con la tendencia al sobreajuste observada anteriormente y refuerza la conclusión de que su estimación de la varianza no ha sido efectiva.

En resumen, el análisis conjunto de las gráficas de predicción, histogramas y residuos confirma que el modelo propuesto no solo se comporta de forma precisa en términos de ajuste cuantitativo, sino que también respeta los principios estadísticos y físicos del problema. Las pérdidas Tobit y CMSE destacan como las más equilibradas al capturar con fidelidad tanto la estructura latente de los datos censurados como la forma general de la respuesta, mientras que las demás alter-

nativas presentan limitaciones más marcadas en uno u otro aspecto, ya sea por no penalizar los errores grandes suficientemente o por una mejorable estimación de la varianza.

### 3.3.3. Evaluación complementaria

Para poner en contexto los resultados obtenidos con redes neuronales, se ha entrenado también un modelo de regresión lineal utilizando exactamente las mismas funciones de pérdida. En todos los casos, el rendimiento fue sustancialmente inferior, tanto en métricas globales como en las zonas más sensibles del dominio, como las regiones censuradas por debajo de 1 MW.

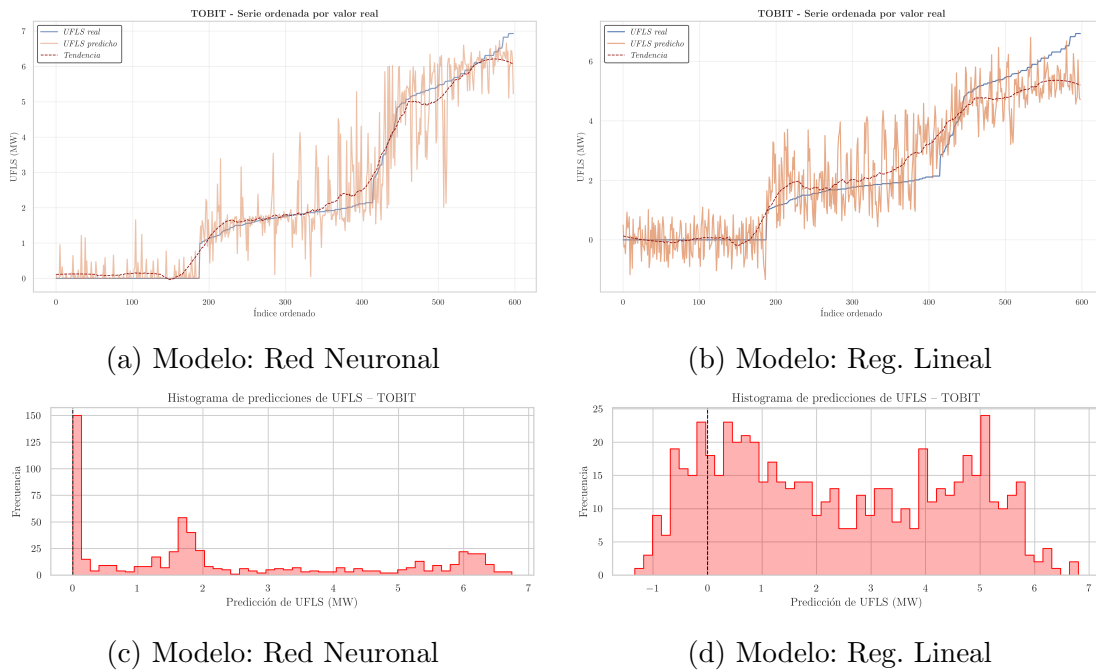


Figura 3.8: Comparativa de la red neuronal frente a la regresión lineal, ambas con pérdida Tobit.

<b>Evaluación</b>	<b>CMAE</b>	<b>CMSE</b>	<b>Tobit</b>	<b>Tobit Hetero</b>
<b>MAE</b>				
Linear Reg.	0.7256	0.7294	0.7282	0.7256
Red Neuronal	0.3662	0.4141	0.3957	0.5446
<b>MSE</b>				
Linear Reg.	0.8755	0.8527	0.8439	0.8503
Red Neuronal	0.5458	0.5063	0.4689	0.6555
<b>NRMSE</b>				
Linear Reg.	0.4056	0.4002	0.3982	0.3997
Red Neuronal	0.3202	0.3084	0.2968	0.3510
<b><math>R^2</math></b>				
Linear Reg.	0.8283	0.8328	0.8345	0.8332
Red Neuronal	0.8929	0.9071	0.9088	0.8714

Cuadro 3.3: Comparativa detallada entre modelos lineales y redes neuronales para cada métrica y función de pérdida.

La Tabla 3.3 resume esta comparativa. Las métricas de error (MAE, MSE, NRMSE) son claramente más altas para el modelo lineal, mientras que el coeficiente de determinación  $R^2$  es notablemente menor en todos los casos, demostrando así un rendimiento bastante inferior por parte del modelo lineal. Esta diferencia no se debe únicamente a la elección de la función de pérdida, sino –sobre todo– a la capacidad de las redes neuronales para capturar relaciones no lineales complejas, que en este problema resultan fundamentales.

La variable objetivo –la potencia deslastrada por la protección– no sigue una relación lineal respecto al estado inicial del sistema, ya que está condicionada por la lógica escalonada del esquema UFLS, la presencia de umbrales y los efectos de la censura. Como consecuencia, los modelos lineales tienden a producir predicciones suavizadas que no reflejan adecuadamente las transiciones abruptas del sistema. Esto se observa claramente en la Figura 3.8, donde se muestra la serie ordenada por valor real: el modelo lineal presenta un ajuste pobre, con tendencia a sobreajustar en regiones bajas y subestimar sistemáticamente en regiones altas, mientras que el modelo neuronal sigue mucho mejor la evolución real con una tendencia suavizada coherente.

Asimismo, los histogramas de predicciones (también en la Figura 3.8) refuerzan esta conclusión. Mientras que el modelo neuronal genera una distribución realista

de valores, respetando la censura inferior pero sin limitarse a replicarla, el modelo lineal distribuye las predicciones de forma casi uniforme sobre el dominio válido, sin reflejar adecuadamente la estructura estadística de los datos. En particular, el modelo lineal no concentra predicciones en torno a cero –como sería esperable por la censura– ni capta correctamente la frecuencia de valores bajos.

Además, es especialmente relevante destacar que, a diferencia de la red neuronal, el modelo lineal produce predicciones negativas en algunos casos, lo cual resulta conceptualmente incorrecto en este contexto como ya se ha explicado.

En conjunto, estos resultados muestran que no basta con utilizar funciones de pérdida adaptadas a datos censurados: su efectividad depende también del modelo con el que se apliquen. Las redes neuronales no solo reducen el error, sino que también reproducen de forma más fiel la lógica subyacente del problema, ofreciendo predicciones coherentes tanto en términos físicos como estadísticos.



# Capítulo 4

## Conclusión

### 4.1. Conclusiones generales del proyecto

El presente proyecto ha abordado el problema de estimar la respuesta de esquemas automáticos de deslastre de carga por baja frecuencia (UFLS) mediante técnicas de *Machine Learning*, centrándose especialmente en el tratamiento de datos censurados. A partir de un conjunto amplio de simulaciones en un entorno eléctrico de alta complejidad, se ha planteado un modelo basado en redes neuronales capaz de predecir la cantidad de carga deslastrada a partir de características del sistema previo al evento.

Uno de los principales aportes del proyecto ha sido la incorporación explícita de la censura en la variable objetivo, un fenómeno fundamental para una estimación realista y robusta. Para ello, se han evaluado múltiples funciones de pérdida, todas ellas adaptadas al marco de datos censurados, entre las que destaca el modelo Tobit [26]. Los resultados empíricos han mostrado que las pérdidas diseñadas específicamente para problemas con censura –en particular la pérdida Tobit clásico– permiten obtener modelos más precisos, estables y coherentes con la estructura estadística del problema.

Otras funciones de pérdida también han mostrado un comportamiento razonable. La pérdida CMSE ha ofrecido un rendimiento competitivo y consistente, destacando por su simplicidad y robustez frente al sobreajuste. La CMAE, por su parte, ha priorizado errores absolutos pequeños y ha resultado útil para minimizar el impacto de valores atípicos, aunque con menor capacidad explicativa global. En cambio, la versión heterocedástica del modelo Tobit no ha alcanzado mejoras sustanciales, mostrando una mayor inestabilidad, probablemente debida a una estimación deficiente de la varianza condicional o a un sobreajuste a ruido estadístico.

En cuanto al modelo propuesto, la red propuesta ha demostrado ser altamente eficaz para capturar las relaciones no lineales que gobiernan la activación del UFLS, superando ampliamente el rendimiento de modelos lineales incluso bajo las mismas funciones de pérdida. Esta diferencia de rendimiento no solo valida el uso de redes neuronales en contextos no lineales con censura, sino que refuerza la idea de que la complejidad estructural del problema requiere arquitecturas capaces de modelar interacciones sutiles y no triviales entre las variables del sistema.

No obstante, uno de los objetivos planteados inicialmente –la posibilidad de trasladar el modelo resultante a una formulación matemática interpretable y aprovechable en esquemas MILP– no ha podido alcanzarse plenamente. Aunque el modelo se inspira conceptualmente en una estructura compatible con la lógica de modelos lineales por partes, su implementación hace uso de funciones de activación suaves como **Softplus**, cuya naturaleza no lineal impide una traducción directa a un sistema de restricciones lineales o enteras. Esta renuncia ha sido necesaria para mantener la capacidad expresiva y predictiva del modelo, dado que el fenómeno estudiado presenta un grado de no linealidad y censura que excede las capacidades de aproximación de formulaciones MILP tradicionales.

Finalmente, este proyecto sienta las bases para futuras investigaciones que busquen un equilibrio entre precisión, interpretabilidad y capacidad de integración en herramientas operativas. Entre las líneas futuras más relevantes se encuentra el diseño de arquitecturas híbridas que combinen módulos neuronales con restricciones explícitas, así como la exploración de modelos generativos adaptados a datos censurados. Otra dirección especialmente prometedora es el desarrollo de modelos cuyas estructuras puedan representarse mediante restricciones lineales o mixtas, lo que permitiría su incorporación directa en formulaciones de optimización energética como MILP. Aunque este objetivo no ha podido alcanzarse plenamente en el presente trabajo –debido al uso de funciones de activación suaves como **Softplus**, incompatibles con esquemas lineales– representa un reto técnico de gran interés para avanzar hacia soluciones más integrables y explicables en contextos de decisión automática.

# Bibliografía

- [1] Andy Hoke et al. “Island Power Systems With High Levels of Inverter-Based Resources: Stability and Reliability Challenges”. En: *IEEE Electrification Magazine* 9.1 (2021), págs. 74-91. DOI: 10.1109/MELE.2020.3047169.
- [2] Md. Nahid Haque Shazon, Nahid-Al-Masood y Atik Jawad. “Frequency control challenges and potential countermeasures in future low-inertia power systems: A review”. En: *Energy Reports* 8 (2022), págs. 6191-6219. ISSN: 2352-4847. DOI: <https://doi.org/10.1016/j.egyrs.2022.04.063>. URL: <https://www.sciencedirect.com/science/article/pii/S2352484722008289>.
- [3] Xuebo Liu et al. “Frequency nadir constrained unit commitment for high renewable penetration island power systems”. En: *IEEE Open Access Journal of Power and Energy* (2024).
- [4] Matteo Fischetti y Jason Jo. “Deep neural networks and mixed integer linear optimization”. En: *Constraints* 23.3 (2018), págs. 296-309.
- [5] Vlad-Rareş Dănăilă y Cătălin Buiu. “A deep learning approach to censored regression”. En: *Pattern Analysis and Applications* 27.1 (2024), pág. 24. ISSN: 1433-755X. DOI: 10.1007/s10044-024-01216-9. URL: <https://doi.org/10.1007/s10044-024-01216-9>.
- [6] Yihui Zuo et al. “Impact of Synchrophasor Estimation Algorithms in ROCOF-Based Under-Frequency Load-Shedding”. En: *IEEE Transactions on Power Systems* 35.2 (2020), págs. 1305-1316. DOI: 10.1109/TPWRS.2019.2936277.
- [7] European Network of Transmission System Operators for Electricity (ENTSO-E). *Technical Background for the Low Frequency Demand Disconnection Requirements*. Inf. téc. Brussels, Belgium: ENTSO-E, 2014.
- [8] Glory Justin y Santiago Paternain. *Data-driven Under Frequency Load Shedding Using Reinforcement Learning*. 2024. arXiv: 2410.04316 [eess.SY]. URL: <https://arxiv.org/abs/2410.04316>.
- [9] Mohammad Rajabdorri et al. *Data-driven Estimation of Under Frequency Load Shedding after Outages in Small Power Systems*. 2024. arXiv: 2312.11389 [eess.SY]. URL: <https://arxiv.org/abs/2312.11389>.

- [10] T Skrjanc, R Mihalic y U Rudez. “A systematic literature review on under-frequency load shedding protection using clustering methods”. En: *Renewable and Sustainable Energy Reviews* 180 (2023), pág. 113294.
- [11] A Darbandsari y T Amraee. “Under frequency load shedding for low inertia grids utilizing smart loads”. En: *International Journal of Electrical Power & Energy Systems* 135 (2022), pág. 107506.
- [12] Turaj Amraee et al. “Probabilistic under frequency load shedding considering RoCoF relays of distributed generators”. En: *IEEE Transactions on Power Systems* 33.4 (2017), págs. 3587-3598.
- [13] JA Laghari et al. “A new under-frequency load shedding technique based on combination of fixed and random priority of loads for smart grid applications”. En: *IEEE Transactions on Power Systems* 30.5 (2014), págs. 2507-2515.
- [14] Enrique Lobato et al. “Robust Optimal Design of UFLS Schemes in Island Power Systems”. En: *Sustainable Energy, Grids and Networks* (2024), pág. 101434.
- [15] Li Zhang y Jin Zhong. “UFLS Design by Using  $f$  and Integrating  $df/dt$ ”. En: *2006 IEEE PES Power Systems Conference and Exposition*. IEEE. 2006, págs. 1840-1844.
- [16] Md Quamrul Ahsan et al. “Technique to develop auto load shedding and islanding scheme to prevent power system blackout”. En: *IEEE transactions on Power Systems* 27.1 (2011), págs. 198-205.
- [17] Zhiqiang Luo et al. “Optimal adaptive decentralized under-frequency load shedding for islanded smart distribution network considering wind power uncertainty”. En: *Applied Energy* 365 (2024), pág. 123162.
- [18] Mingyu Sun et al. “Underfrequency load shedding using locally estimated RoCoF of the center of inertia”. En: *IEEE Transactions on Power Systems* 36.5 (2021), págs. 4212-4222.
- [19] Tomislav Baškarad et al. “Dynamically adaptive method for under frequency load shedding protection scheme reconfiguration”. En: *Electric power systems research* 207 (2022), pág. 107823.
- [20] Bei Xu et al. “Under-Frequency Load Shedding for Power Reserve Management in Islanded Microgrids”. En: *IEEE Transactions on Smart Grid* (2024).
- [21] Yichen Zhang et al. “Encoding Frequency Constraints in Preventive Unit Commitment Using Deep Learning With Region-of-Interest Active Sampling”. En: *IEEE Transactions on Power Systems* 37.3 (2022), págs. 1942-1955. DOI: 10.1109/TPWRS.2021.3110881.

- 
- [22] Chao-Yuan Lai y Chih-Wen Liu. “A Scheme to Mitigate Generation Trip Events by Ancillary Services Considering Minimal Actions of UFLS”. En: *IEEE Transactions on Power Systems* 35.6 (2020), págs. 4815-4823. DOI: 10.1109/TPWRS.2020.2993449.
- [23] Mohammad Rajabdorri, Enrique Lobato y Lukas Sigríst. “Robust frequency constrained uc using data driven logistic regression for island power systems”. En: *IET Generation, Transmission & Distribution* 16.24 (2022), págs. 5069-5083. DOI: <https://doi.org/10.1049/gtd2.12658>. eprint: <https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/gtd2.12658>. URL: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/gtd2.12658>.
- [24] Fei Teng y Goran Strbac. “Full Stochastic Scheduling for Low-Carbon Electricity Systems”. En: *IEEE Transactions on Automation Science and Engineering* 14.2 (2017), págs. 461-470. DOI: 10.1109/TASE.2016.2629479.
- [25] Mohammad Rajabdorri et al. “Data-driven estimation of the amount of under frequency load shedding in small power systems”. En: *Engineering Applications of Artificial Intelligence* 139 (2025), pág. 109617.
- [26] James Tobin. “Estimation of Relationships for Limited Dependent Variables”. En: *Econometrica* 26.1 (1958), págs. 24-36. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/1907382> (visitado 30-06-2025).
- [27] Bhairavi Premnath y Anastasia Sofroniou. “Analysing Load Shedding to Increase Stability in the Swing Equation”. En: *Mathematics* 13.8 (2025). ISSN: 2227-7390. DOI: 10.3390/math13081314. URL: <https://www.mdpi.com/2227-7390/13/8/1314>.
- [28] Renhai Feng et al. “Adaptive non-parametric kernel density estimation for under-frequency load shedding with electric vehicles and renewable power uncertainty”. En: *Scientific Reports* 15.1 (abr. de 2025), pág. 11499. ISSN: 2045-2322. DOI: 10.1038/s41598-025-94419-x. URL: <https://doi.org/10.1038/s41598-025-94419-x>.

## *BIBLIOGRAFÍA*

---