Trabajo Fin de Máster

Máster Universitario en Ingeniería Aeronáutica

Diseño de un Sistema tolerante a Ciberataques mediante Control Predictivo Basado en Modelo

Autor: Simón Gutiérrez de Ravé Serrano

Tutora: Ascensión Zafra Cabeza

Dpto. de Ingeniería de Sistemas y Automática
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2024

Trabajo Fin de Máster
Máster Universitario en Ingeniería Aeronáutica

Diseño de un Sistema tolerante a Ciberataques mediante Control Predictivo Basado en Modelo

Autor:

Simón Gutiérrez de Ravé Serrano

Tutora:

Ascensión Zafra Cabeza

Profesora titular

Dpto. de Ingeniería de Sistemas y Automática

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2024

Trabajo de Fin de Máster: Diseño de un Sistema tolerante a Ciberataques mediante Control Predictivo Basado en Modelo

Autor: Simón Gutiérrez de Ravé Serrano

Tutora: Ascensión Zafra Cabeza

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

El Secretario del Tribunal

Fecha:

Resumen

La ciberseguridad se presenta como uno de los aspectos más importantes en la mayoría de los procesos informáticos de la actualidad, siendo la variedad de los ciberataques cada vez más extensa y compleja. Es por ello por lo que, la investigación y desarrollo de programas y algoritmos los cuales sean capaces de detectar estos ataques se ha convertido en un pilar fundamental y necesario para velar por la seguridad de todo aquel equipo, red o sistema que pueda ser objeto de este tipo de delincuencia.

El objetivo de este proyecto es la simulación de ciberataques y comprobación del funcionamiento de métodos de detección que puedan ser eficaces para este tipo particular de actos ilícitos. Para ello se trabajará sobre una microrred experimental, la cual se encuentra ubicada en el Centro de Experimentación de El Arenosillo (CEDEA),[1].

A lo largo del documento se desarrollará toda la información necesaria para comprender el funcionamiento de una microrred y los conceptos básicos de los ciberataques. Posteriormente se detallarán los fundamentos teóricos de los métodos de detección y finalmente se realizarán las simulaciones de los ciberataques a distintos elementos de la microrred.

Abstract

Cybersecurity emerges as one of the most critical aspects in the majority of current computer processes, with the variety of cyber attacks becoming increasingly extensive and complex. This is why the research and development of programs and algorithms capable of detecting these attacks have become a fundamental and necessary pillar to ensure the security of any equipment, network, or system that may be the target of this type of crime.

The objective of this project is the simulation of cyber attacks and the verification of the functionality of detection methods that can be effective for this particular type of illicit activity. To achieve this, we will work on an experimental microgrid, which is located at the El Arenosillo Experimentation Center (CEDEA),[1].

Throughout the document, all the necessary information will be developed to understand the operation of a microgrid and the basics of cyber attacks. Subsequently, the theoretical foundations of detection methods will be detailed, and finally, simulations of cyber attacks on different elements of the microgrid will be carried out.

Índice

Notación

$x(t)$	Valor de x en el instante de t
$x(t+k\|t)$	Valor de x en el instante $t+k$ calculado en el instante t
$\hat{x}$	Valor estimado de x
$\tilde{x}$	Error en la estimación de x
A^T	Traspuesta de A
A^{-1}	Inversa de la matriz A
$e.o.c.$	En cualquier otro caso
$P(A)$	Probabilidad del suceso A
$E[X]$	Valor esperado de la variable aleatoria X
σ_X	Desviación estándar de la variable aleatoria X
$F(x)$	Función de Distribución
F^{-1}	Inversa de la Función de Distribución
$f(x)$	Función de Densidad
$P(t)$	Matriz de covarianza del error de una estimación en el instante t
I_n	Matriz identidad de dimensión n
A_{mxn}	Matriz A con m filas y n columnas
=	Igual que
$<$	Menor que
$>$	Mayor que
$\leq$	Menor o igual
$\geq$	Mayor o igual
x	Vector de estados

u	Vector de entradas manipulables
y	Vector salidas
d	Vector de perturbaciones
T_s	Tiempo de muestreo
$\sum_{i=1}^{N} x$	Sumatorio de $i = 1$ hasta N de x
$máx\ x(t)$	Máximo de $x(t)$
$mín\ x(t)$	Mínimo de $x(t)$
$s.t.$	Sujeto a
$b \cdot 10^{a}$	Formato científico

1 Introducción

1.1 Situación actual

En la era digital actual, nos encontramos sumergidos en una constante conectividad y dependencia de las conocidas como tecnologías de la información y la comunicación (TIC). La abundancia de dispositivos y elementos interconectados ha dado lugar a la formación de microrredes, entornos de sistemas distribuidos formados por generadores y cargas de modo que sea posible abastecerse de manera autónoma mediante la combinación de sistemas de generación tradicionales y de origen renovable, de manera que se gestione la demanda del modo más eficiente posible con vistas a prolongar la vida útil de los componentes, así como un ahorro energético [2].

La utilización y creación de las microrredes hubiese sido inviable años atrás dado el elevado coste de sus componentes y la baja eficiencia de estos, sin embargo, las recientes mejoras en estos elementos han contribuido a que la creación de microrredes sea una realidad viable gracias a:

- Sus beneficios en términos de sostenibilidad medioambiental
- La caída de los costes de las tecnologías de almacenamiento energético y sistemas de generación de energías renovables (paneles solares, aerogeneradores…)
- El desarrollo de sistemas de control inteligente que convierten a las microrredes en entornos inteligentes permitiendo la gestión activa de recargas eléctricas y energía almacenada, reduciendo los costes del suministro eléctrico.

El presente trabajo aborda el desafío crítico de la detección de ciberataques que pudieran ocurrir en estas microrredes, para ello se utilizarán algoritmos basados en modelos, los cuales se desarrollarán y explicarán a lo largo de la elaboración de este documento. La seguridad cibernética se ha convertido en una prioridad para la gran mayoría de las empresas hoy día, así como para los usuarios diarios de internet, ya que la elaboración y diversificación de los ataques cibernéticos han ido evolucionando de manera exponencial. No son una excepción las microrredes, con una compleja infraestructura y variedad de dispositivos interconectados, las cuales pueden también sufrir de estos ciberataques, siendo una gran preocupación para los usuarios de estas.

Este trabajo se centra en la detección de ciberataques en microrredes, combinando la teoría recopilada en tema de ciberseguridad con las capacidades de los algoritmos basados en modelo con la esperanza de fortalecer la resiliencia de las microrredes en un entorno digital cada vez más afectado por los ataques cibernéticos.

1.2 Aplicación práctica del proyecto

Una vez introducido el tema principal a abordar, la mejor manera de entender cómo sería la situación real/práctica a la que nos enfrentamos es buscando una aplicación directa de la problemática a tratar.

Para ello, se procede a realizar la simulación de los hipotéticos ciberataques sobre una microrred experimental situada en el "Centro de Experimentación De El Arenosillo "(CEDEA), ubicado en Huelva y perteneciente al Instituto Nacional de Técnica Aeroespacial (INTA).

La microrred por estudiar está compuesta por un conjunto de placas fotovoltaicas, un aerogenerador, baterías de ácido plomo-litio y de ion litio, un electrolizador, una carga y una fuente de alimentación programables, supercondensadores y un punto de recarga para vehículos híbridos y uno eléctrico.

Está controlada por un controlador predictivo basado en modelo (*Model Predictive Control,* MPC*)* ([3],[4],[5]), el cual se encarga de controlar los intercambios de potencia entre los elementos que conforman la microrred.

El MPC, es un tipo de control clasificado como técnica de control óptimo, basado en un modelo lineal del sistema que calcula señales de control óptimas para enviar al sistema a partir de un estado estimado a lo largo de un horizonte y de las señales de control anteriormente enviadas.

1.3 Ejecución

Para el desarrollo de las pruebas y simulaciones se ha partido de un modelo implementado en Simulink®, un entorno de programación visual que funciona sobre el entorno de programación Matlab® (MATri LABoratory) [6] Simulink® es una herramienta de simulación de modelos o sistemas, la cual permite modelar la microrred gráficamente mediante bloques e introducir funciones de Matlab® que se ejecutan en combinación con la simulación del modelo creado en Simulink®.

Sobre el modelo original, se realizarán las modificaciones pertinentes, con el objetivo de que las simulaciones realizadas se correspondan con posibles ciberataques a la microrred.

2 Microrredes: Funcionamiento y Tipología

En la situación actual de búsqueda constante de soluciones sostenibles, las microrredes emergen como elementos clave en la transformación de energía distribuida. Este capítulo se adentra en el ecosistema de las microrredes, tratando de explorar los mecanismos de funcionamiento y tipologías de estas microrredes, las cuales abarcan desde aplicaciones residenciales hasta entornos industriales.

El rasgo que distingue a las microrredes es su capacidad para operar de manera autónoma o conectada a la red principal, adaptándose a las demandas específicas de comunidades, instalaciones comerciales y sectores industriales (Figura 2.1).

En resumen, en este capítulo se proporcionará una visión general de las microrredes, desde sus fundamentos técnicos hasta su aplicación práctica.

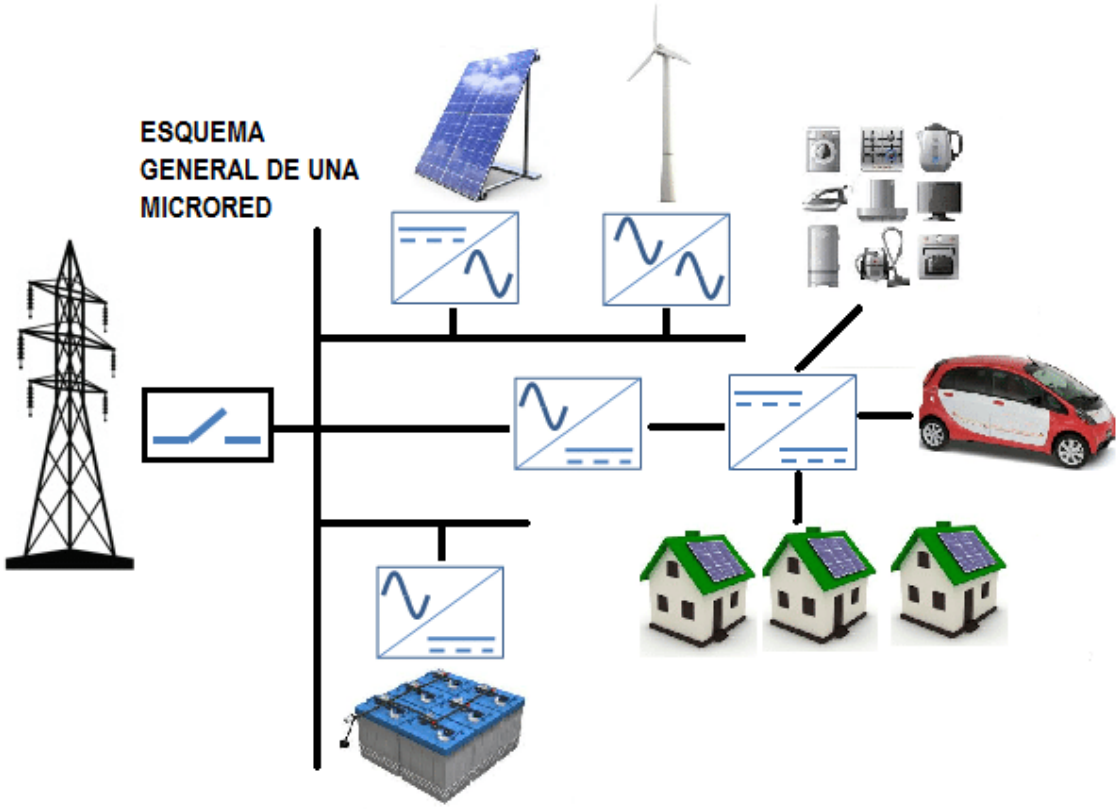

Figura 2.1 *Modelo general de una Microrred. Fuente [7]*

2.1 Ecosistema de las Microrredes

2.1.1 Definición y características

Las microrredes se definen como sistemas locales de generación, almacenamiento y consumo de electricidad. A diferencia de las redes eléctricas convencionales, las microrredes tienen capacidad de operar de manera autónoma o conectadas a la red principal, aportando flexibilidad y resiliencia ante posibles interrupciones. De entre las características particulares de las microrredes se incluyen la diversificación de fuentes de energía, la integración de tecnologías de almacenamiento (baterías) y sistemas de gestión avanzada.

2.1.2 Importancia en la transición energética

En la transición energética las microrredes desempeñan un papel crucial. Su capacidad para incorporar fuentes de energía renovable, como solar o eólica, contribuye significativamente a la reducción de emisiones de carbono y a la creación de sistemas más limpios y eficientes. Además, las microrredes permiten la descentralización de la generación de energía, promoviendo la autonomía local y disminuyendo la dependencia de fuentes centralizadas.

2.1.3 Componentes principales

El ecosistema de las microrredes comprende varios componentes interconectados que colaboran para garantizar su funcionamiento de manera eficiente. Entre los elementos clave se encuentran [8]:

- **Generación descentralizada:** Incorporando diversas fuentes de energía, como paneles solares, aerogeneradores o sistemas de cogeneración, permitiendo una producción adaptada a las necesidades específicas de la microrred.
- **Almacenamiento de energía:** Utilizando tecnologías de almacenamiento como baterías o sistemas de almacenamiento térmico que sean capaces de gestionar el excedente de energía durante periodos de baja demanda y liberarlo cuando la demanda es alta.
- **Gestión inteligente:** Implementando sistemas de control avanzados que monitorizan la demanda, ajustan la generación en tiempo real y optimizan el uso de recursos disponibles.
- **Redes de comunicación:** Facilitando la comunicación entre las diferentes partes de una microrred, permitiendo una coordinación eficaz y la transmisión de datos esenciales para la gestión y control.

Estos componentes otorgan la capacidad de adaptarse a variaciones en la demanda, cambios en las condiciones climáticas y eventos imprevistos, consolidando así la idea de que las microrredes son una solución integral en la búsqueda de sistemas energéticos más sostenibles y resilientes.

2.2 Funcionamiento de las Microrredes

2.2.1 Operación autónoma frente conexión a la red principal

La característica principal de las microrredes es su capacidad para operar tanto de manera autónoma como conectada a la red principal [2]. En el modo autónomo, la microrred es capaz de generar, almacenar y distribuir energía de manera independiente, lo que resulta de extrema utilidad en situaciones de emergencia o en ubicaciones muy remotas en las cuales la conexión a la red principal no sea posible o sea limitada. Sin embargo, la operación de la microrred funcionando conectada a la red principal permite obtener el beneficio de la estabilidad y los recursos adicionales proporcionados por la red global, contribuyendo a la eficiencia y flexibilidad del sistema.

2.2.2 Integración de fuentes de energía renovable

Las fuentes de energía renovable (Figura 2.2) constituyen un pilar fundamental en el funcionamiento de las microrredes. La diversificación de estas fuentes, como la energía solar y eólica, no solo reduce la dependencia a los combustibles fósiles, sino que también promueve el uso de una energía más limpia y sostenible. La capacidad de adaptar la generación a las condiciones climáticas locales permite conseguir una producción eficiente.

Figura 2.2: *Fuentes de energía renovables. Fuente: [9]*

2.2.3 Gestión inteligente de recursos

La gestión inteligente es fundamental para maximizar la eficiencia y la resiliencia de las microrredes. Sistemas de control avanzados supervisan continuamente la demanda de energía, la disponibilidad de recursos y las condiciones operativas [10]. Mediante los datos extraídos por esos sistemas de control, la microrred ajusta dinámicamente la generación y el almacenamiento para satisfacer la demanda en tiempo real. La implementación de algoritmos de optimización y aprendizaje automático facilita a la microrred la toma de decisiones autónomas y la adaptación a patrones de consumo variados, mejorando la eficiencia operativa a lo largo del tiempo, dadas unas condiciones cambiantes según los periodos estacionales del año.

2.3 Tecnologías subyacentes

2.3.1 Sistemas de almacenamiento de energía

Los sistemas de almacenamiento de energía son un elemento fundamental para la operación eficiente de las microrredes. Al integrar diferentes tecnologías de almacenamiento, como baterías de iones de litio, sistemas de almacenamiento térmico o tecnologías de almacenamiento de energía mecánica, las microrredes pueden gestionar eficazmente la variabilidad de la generación y la demanda (Figura 2.3). Los sistemas de almacenamiento de energía permiten almacenar la energía en los periodos en los que la demanda es baja con la posibilidad de liberar la energía durante momentos de alta demanda, mejorando así la estabilidad y confiabilidad del suministro [11].

Figura 2.3: *Sistema de baterías en una microrred. Fuente: [11]*

2.3.2 Redes de comunicación en Microrredes

Las redes de comunicación son el eje principal que permite la coordinación y el intercambio de información entre los diversos componentes de las microrredes. Utilizando protocolos de comunicación avanzados tales como el internet de las cosas en inglés “Internet of Things” (IoT) y las tecnologías de red inalámbrica, las microrredes pueden recopilar datos en tiempo real sobre la generación, el almacenamiento y la demanda de energía. Esto facilita la toma de decisiones autónoma y la optimización del rendimiento de la microrred [12].

2.3.3 Automatización y control

La automatización y el control avanzado son esenciales para la gestión eficiente de las microrredes. Mediante sistemas de control centralizados o distribuidos que utilicen algoritmos inteligentes, se consigue monitorear y regular tanto la generación de energía como su consumo. La automatización permite una respuesta rápida a condiciones operativas, optimizando continuamente la condición de la microrred. Con la combinación de estas dos tecnologías se consigue que las microrredes puedan operar de manera eficiente y adaptarse a las dinámicas cambiantes del entorno [13].

2.4 Tipos de Microrredes

En el año 2021, se crea la organización “Think Microgrid” la cual tiene como finalidad ser la voz unificada de la industria de las microrredes, destacando el papel que desempeñarán las microrredes en un momento único de la industria [14].

En la página web de la organización se encuentra la siguiente definición de Microrred:

“Una microrred es un sistema de energía autosuficiente que sirve a una huella geográfica puntual, como un campus universitario, un complejo hospitalario, un centro de negocios o un vecindario”.

Como se puede apreciar el del concepto de microrred es algo muy variado y se pueden encontrar diversas definiciones según la fuente consultada, aun así, la organización “Think Microgrid” realiza una amplia distinción en la tipología de estas microrredes por lo que se ha considerado una buena referencia para hacer ver al lector la gran variedad de casos de microrred que existen actualmente.

Se tienen entonces los siguientes tipos de microrred:

- **Avanzada:** Microrred que utiliza un software sofisticado, controles e inteligencia artificial para gestionar múltiples recursos energéticos. La mayoría de los tipos de microrredes enumerados a continuación podrían configurarse como microrredes avanzadas.

- **Microrred universitaria:** Una microrred que da servicio a varios edificios en una única parcela grande de terreno tal como un centro médico, parque empresarial o centro educativo.
- **Microrred comunitaria:** Microrred que da servicio a instalaciones críticas dentro de una comunidad, como centros de respuesta a emergencias, plantas de tratamiento de agua y aguas residuales, tiendas de comestibles, estaciones de servicio, edificios gubernamentales y refugios. También podrían extenderse a edificios u hogares no críticos.
- **Microrred conectada a la red:** Tal y como su nombre indica, es una microrred que está conectada a la red eléctrica central, pudiendo separarse de esta cuando las condiciones sean oportunas.
- **Microrred híbrida:** Pueden generar energía con dos o más fuentes de energía distribuidas, como pueden ser la eólica y solar. Pueden funcionar tanto de manera autónoma como conectadas a la red.
- **Microrred móvil:** Las microrredes móviles se pueden reubicar para apoyar a los equipos de respuesta a emergencias o proporcionar energía para cargar dispositivos electrónicos o médicos en situaciones críticas. Debido al uso frecuente de energías renovables o baterías, las microrredes móviles reducen la necesidad de enviar combustible a áreas remotas o zonas de emergencia. Un ejemplo de esta microrred móvil sería aquella que implique el uso de vehículos eléctricos para respaldar la red en momentos de alta demanda.
- **Nanorred:** Son básicamente pequeñas microrredes que sirven a un solo cliente o instalación. Estas microrredes normalmente pueden funcionar tanto en modo conectado a la red como en modo autónomo.
- **Microrred remota:** Las microrredes remotas se encuentran en islas, áreas aisladas o diferentes lugares en los cuales no hay proximidad de una red eléctrica confiable, por lo que no tienen posibilidad de estar conectadas a la red.
- **Microrred de energía renovable:** Tal y como su nombre indica es una microrred que utiliza una o más fuentes de energía renovables.
- **Microrred de servicios públicos:** Una microrred que pertenece y es operada por una empresa de servicios públicos o propiedad de un inversor.

2.5 Desafíos, oportunidades y líneas futuras

2.5.1 Desafíos

A pesar del notable beneficios de la utilización de las microrredes, su implementación no está exenta de retos. El cumplimiento de la normativa y aspectos regulatorios pueden presentarse como barreras para su adopción en la sociedad, por lo que esta necesidad de crear estándares comunes y marcos regulatorios se destaca como un desafío clave.

Como reto particular con relación al desarrollo de este proyecto, se tiene la gestión de la ciberseguridad en entornos distribuidos, la cual, plantea preocupaciones críticas, exigiendo estrategias robustas para proteger las microrredes de posibles amenazas.

2.5.2 Oportunidades

En medio de los desafíos antes descritos, se presentan numerosas oportunidades en el desarrollo de estas microrredes, ya que suponen la vía para una transición energética sostenible, fomentando la investigación y desarrollo de tecnologías cada vez más eficientes.

Son también una gran oportunidad de negocio innovador debido a su flexibilidad y adaptabilidad, promoviendo la participación de la comunidad y la colaboración entre diferentes actores del sector energético.

2.5.3 Líneas futuras

La evolución de las microrredes promete una serie de desarrollos muy optimistas. A medida que la tecnología va avanzando, se espera que las microrredes se vuelvan aún más inteligentes y autónomas. Gracias a la integración de tecnologías cada vez más asentadas, como son la inteligencia artificial y el aprendizaje automático, se conseguirá una gran mejoría en la toma de decisiones mejorando la eficiencia operativa y la capacidad de respuesta a cambios en las condiciones del entorno [15].

El desarrollo e investigación en materia de almacenamiento de energía también desempeñará un papel fundamental en el futuro de las microrredes. Se espera conseguir avances en baterías de mayor capacidad, sistemas de almacenamiento más eficientes y nuevas formas tanto de almacenar como de utilizar esta energía que en su conjunto llevarán a una gestión más efectiva de los recursos energéticos.

Como conclusión, las líneas futuras de las microrredes apuntan hacia una mayor inteligencia, eficiencia y colaboración con el sector energético, lo cual está en consonancia con la problemática actual energética que busca constantemente una transición hacia una mejor gestión energética y sostenibilidad.

3 Teoría y Conocimientos Generales de Ciberataques

El mundo de los ciberataques es extremadamente amplio y difícil de abarcar en su totalidad para un proyecto como este, el cual, no trata este tema en su totalidad, sino que lo aplica ligado a otro concepto igualmente amplio como es la microrred.

Es por ello por lo que en este capítulo se tratará de transmitir al lector la información necesaria para comprender el concepto de ciberataque de la manera más compacta posible.

3.1 Fundamentos de ciberseguridad

3.1.1 Definición y evolución de ciberataques

La ciberseguridad se sitúa en el epicentro de la protección contra amenazas digitales en constante evolución. Se define como la práctica de proteger sistemas informáticos, redes y programas contra ataques no autorizados, daños o acceso no deseado [16]. A lo largo del tiempo, la evolución de los ciberataques ha sido notable, pasando de intrusiones básicas a sofisticadas campañas llevadas a cabo por actores estatales y no estatales. Comprender la evolución de estas tácticas es crucial para diseñar estrategias efectivas de defensa cibernética.

Como dato curioso en cuanto al origen o posible origen de los ciberataques, el primer "ciberataque" tuvo lugar en el año 1834 en Francia [17]. El francés Claude Chappe (Figura 3.1), inventor del telégrafo óptico, fue capaz de crear un sistema nacional que podía transmitir información la cual se utilizaba para transmitir el valor de los bonos en los movimientos de los mercados. Dos banqueros franceses, François y Joseph Blanc, encontraron la forma de utilizar el telégrafo para sus propios fines y conseguir conocer el valor de los bonos antes que sus competidores. Buscando la manera perfecta de conseguir la información antes que sus competidores, se les ocurrió la idea de sobornar al operador de telégrafos ópticos con una buena suma de dinero, a cambio de que introdujera información oculta sobre el valor de los bonos en los mensajes del Gobierno. El truco, en realidad, fue incluir símbolos de retrocesos justo después del mensaje sobre la situación del mercado. Dicho símbolo indicaba que el carácter anterior debía ser ignorado. El dato oculto incluido por el operador en cada mensaje era interpretado por un ayudante que se lo comunicaba directamente a los hermanos Blanc. El "hackeo" fue descubierto 2 años después, en 1836, cuando el operario sobornado se puso enfermo y su substituto develó la estafa. No obstante, los hermanos Blanc no fueron condenados puesto que por aquel entonces no había ninguna ley que prohibiera este tipo de acciones.

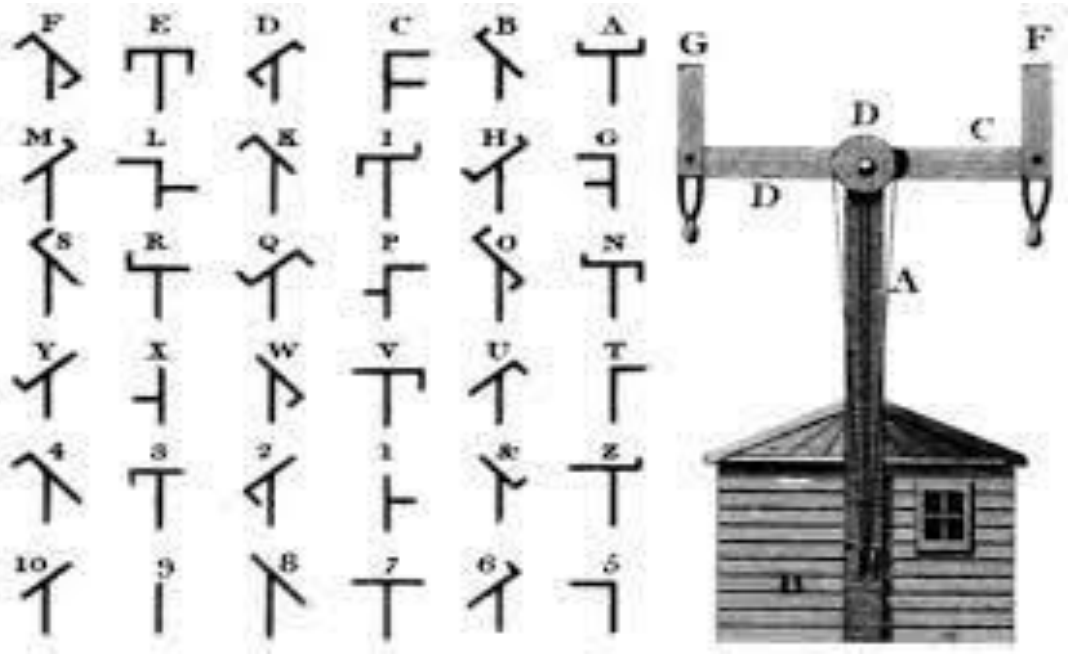

Figura 3.1: *Sistema de telégrafo óptico inventado por Claude Chappe. Fuente: [17]*

La anécdota anterior sirve de ejemplo para ver que la intención de manipular y atacar sistemas informáticos es tan remota como el propio origen de estos sistemas. Es por ello por lo que la importancia de protegerse ante estos intentos de ataque es de vital relevancia.

3.1.2 Amenazas y actores en ciberseguridad

Las amenazas en los ciberataques son muy variadas y van desde individuos malintencionados hasta grupos organizados con motivaciones financieras, políticas o simplemente destructivas. Los actores pueden incluir desde ciberdelincuentes individuales hasta hackers respaldados por estados nacionales. Es fundamental reconocer la diversidad de estas amenazas y entender las motivaciones subyacentes para implementar medidas de seguridad proporcionadas y adaptativas.

3.1.3 Principales objetivos de los ciberataques

Los ciberataques pueden tener una gran variedad de objetivos, desde el robo de información confidencial y la interrupción del funcionamiento de los sistemas hasta la manipulación de algoritmos para obtener beneficios financieros. El conocimiento de estos objetivos es fundamental para anticipar y mitigar los posibles ciberataques.

La pérdida de datos, la interrupción del servicio y la degradación de la confianza del usuario son algunas de las posibles consecuencias potenciales de los ciberataques, confirmando así la necesidad de una ciberseguridad sólida en todas las capas de la infraestructura digital [18].

3.2 Tipos de ciberataques

A continuación, se procederá a describir los tipos de ciberataques más comunes [19], como aspecto principal a destacar de todos ellos es que, a pesar de las diferencias que hay en su funcionamiento, no son independientes unos de otros. En numerosas ocasiones un ataque completo está compuestos por varias de estas técnicas.

3.2.1 Ataques de denegación de servicios (DDoS)

Los ataques de denegación de servicio buscan colapsar un sistema, red o servicio con tráfico malicioso, provocando su saturación y la imposibilidad de atender solicitudes legítimas. En el contexto de las microrredes, un ataque DDoS podría paralizar la capacidad de la red para operar normalmente, afectando la disponibilidad y confiabilidad del suministro energético.

3.2.2 Malware y Ransomware

El malware (programa maligno) es un software malicioso diseñado para dañar o infiltrarse en sistemas, el cual puede ser una amenaza significativa para las microrredes.

El ransomware (secuestro y rescate de datos) es una forma particular del malware, cifra datos críticos y exige un rescate para su liberación.

Enfocado a las microrredes, el malware podría comprometer la integridad de los sistemas de control y gestión de las microrredes, mientras que el ransomware podría paralizar completamente las operaciones.

3.2.3 Ingeniería social y Phishing

La ingeniería social implica manipular a individuos para obtener información confidencial. El phishing, una forma común de ingeniería social, utiliza tácticas engañosas para que los usuarios revelen datos personales o credenciales de acceso. Estos ataques pueden ser instrumentos de entrada tratando de hacer caer en la trampa a las personas al cargo de los sistemas informáticos de una microrred, comprometiendo así la seguridad de la microrred.

3.2.4 Ataques de fuerza bruta

Los ataques de fuerza bruta intentan descifrar contraseñas o claves de acceso probando sistemáticamente todas las combinaciones posibles (de ahí el nombre de fuerza bruta).

En el contexto de las microrredes, un ataque de fuerza bruta exitoso podría comprometer la seguridad de los sistemas de control, provocando la extracción de credenciales de seguridad, con todas las posibles consecuencias que pudiese haber en la integridad y confidencialidad de la red.

3.2.5 Gusanos, virus y troyanos

Los gusanos, son ciberataques que tienen la capacidad de propagarse dentro de una misma red por ellos mismos sin la intervención del usuario.

Al contrario que los gusanos, los virus necesitan de la intervención de un usuario para propagarse. Los virus suelen alojarse en archivos o programas que al ejecutarse afectan al terminal.

A diferencia de los dos anteriores, los troyanos no se propagan y se camuflan como una aplicación legítima, aunque también necesitan ser ejecutados para liberarse. Su diferencia principal es que abren una puerta trasera o "backdoor" que permite acceder al sistema.

3.2.6 Spyware y Adware

Los spyware son programas espía, que se ejecutan automáticamente cada vez que se enciende un equipo infectado y son capaces de recabar información que posteriormente enviarán a un tercero sin la autorización del usuario.

Los adwares, son softwares basados en publicidad, los cuales muestran anuncios en forma de ventanas emergentes. Aunque puedan resultar muy molestos, no son maliciosos por sí mismos, pero sí que pueden llevar algún tipo de spyware implícito.

3.3 Ciberseguridad en microrredes

3.3.1 Importancia de la ciberseguridad en microrredes

La importancia de la ciberseguridad en el contexto de las microrredes es fundamental dada la creciente interconexión y dependencia a las tecnologías digitales. La robustez y fiabilidad de las microrredes, que desempeñan un papel fundamental en la gestión de la energía descentralizada, dependen en gran medida de la integridad de los sistemas informáticos subyacente [10].

La ciberseguridad se convierte entonces en un pilar crítico para garantizar el funcionamiento seguro y continuo de las microrredes, protegiéndolas contra posibles amenazas y ataques maliciosos.

A continuación, se mencionan algunos aspectos que justifican la importancia de la ciberseguridad en las microrredes [20]:

- **Resiliencia frente a ataques:** Las microrredes son sistemas descentralizados que integran fuentes de energía distribuidas, como paneles solares, turbinas eólicas, baterías, etc. Esto las hace potencialmente vulnerables a ataques cibernéticos que podrían comprometer su funcionamiento. Garantizar la seguridad de estos sistemas es esencial para evitar interrupciones en el suministro de energía y para mantener la continuidad de servicio.

- **Protección de datos sensibles:** Las microrredes pueden involucrar la monitorización y gestión de datos sensibles, como información sobre el consumo de energía, patrones de uso, datos financieros, entre otros. Es crucial implementar medidas de seguridad robustas para proteger estos datos y prevenir accesos no autorizados que puedan comprometer la privacidad de los usuarios y la integridad de la red.
- **Prevención de intrusiones físicas y cibernéticas:** Las microrredes pueden estar compuestas por una combinación de sistemas físicos y digitales. La seguridad no solo se refiere a la protección de los sistemas informáticos y de comunicación, sino también a la seguridad física de los componentes de la red. Esto implica implementar medidas para prevenir el acceso no autorizado a infraestructuras críticas, como subestaciones eléctricas, centros de control, etc.
- **Integración con la red principal:** En muchos casos, las microrredes están conectadas a la red eléctrica principal. Esta conexión introduce posibles puntos de vulnerabilidad que podrían ser explotados por atacantes externos. Es crucial implementar mecanismos de autenticación y cifrado para garantizar la seguridad de las comunicaciones entre la microrred y la red principal, así como para proteger contra posibles intrusiones desde la red principal hacia la microrred.
- **Mantenimiento de la confiabilidad y calidad del suministro eléctrico:** La seguridad en las microrredes también está estrechamente relacionada con la confiabilidad y calidad del suministro eléctrico. Los ataques cibernéticos o físicos pueden afectar la operación normal de la microrred y causar interrupciones en el suministro eléctrico, lo que a su vez puede tener consecuencias económicas y sociales significativas.

En resumen, la seguridad en las microrredes es esencial para garantizar la confiabilidad, disponibilidad e integridad de los sistemas de energía distribuida, así como para proteger la privacidad y los datos sensibles de los usuarios. Se requiere una combinación de medidas técnicas, operativas y de gestión de riesgos para mitigar las amenazas y asegurar el funcionamiento seguro de las microrredes.

3.3.2 Vulnerabilidades específicas de las microrredes

Las microrredes presentan vulnerabilidades particulares que requieren especial atención en términos de ciberseguridad. Debido a la diversidad de componentes, desde generadores distribuidos hasta sistemas de almacenamiento y dispositivos de gestión, cada uno de ellos representa un posible punto de entrada para un atacante. La interconexión de estos elementos, si no se gestiona adecuadamente, puede amplificar el impacto de los ataques. Además, la operación autónoma de algunas microrredes puede introducir desafíos adicionales al limitar la supervisión directa.

Algunas de las vulnerabilidades más relevantes son las siguientes [21]:

- **Interconexión de sistemas heterogéneos:** Las microrredes suelen integrar una variedad de sistemas heterogéneos, como generadores de energía renovable, sistemas de almacenamiento de energía, dispositivos de medición inteligente, sistemas de control de red, etc. Esta diversidad de sistemas puede generar una

superficie de ataque más amplia y compleja, ya que cada componente puede tener diferentes protocolos de comunicación y niveles de seguridad.

- **Comunicaciones inalámbricas y protocolos expuestos:** Muchas microrredes utilizan comunicaciones inalámbricas para la monitorización y control de los dispositivos, lo que puede aumentar la exposición a ataques cibernéticos, como el acceso no autorizado, la interceptación de datos y la suplantación de identidad.
- **Gestión remota y accesibilidad:** La capacidad de gestionar y supervisar las microrredes de forma remota puede ser una ventaja en términos de eficiencia operativa, pero también puede introducir riesgos de seguridad significativos. Los sistemas de gestión remota pueden ser blanco de ataques de denegación de servicio (DoS), intrusiones y explotación de vulnerabilidades si no se implementan adecuadas medidas de autenticación, autorización y cifrado.
- **Dependencia de la red eléctrica principal:** Las microrredes conectadas a la red eléctrica principal pueden estar expuestas a vulnerabilidades inherentes de esa red, como ataques cibernéticos dirigidos a la infraestructura de transmisión y distribución, así como a problemas de calidad de energía que podrían ser propagados a la microrred.
- **Falta de actualizaciones de seguridad:** Algunas microrredes pueden estar compuestas por equipos y sistemas más antiguos que no reciben actualizaciones de seguridad regulares o no son compatibles con las últimas medidas de protección cibernética. Esto deja a estas microrredes vulnerables a atawues conocidos y ataques dirigidos.
- **Factores humanos:** Las vulnerabilidades en las microrredes también pueden surgir debido a errores humanos, como la falta de conciencia sobre la seguridad cibernética, contraseñas débiles, acceso no autorizado de personal no autorizado, entre otros.

Las vulnerabilidades anteriormente citadas suponen una serie de desafíos específicos en términos de ciberseguridad debido a su naturaleza descentralizada, diversidad de sistemas, dependencia de la red principal y necesidad de gestión remota. Abordar estas vulnerabilidades requiere un enfoque integral que incluya la implementación de medidas de seguridad técnicas, operativas y de gestión de riesgos.

3.3.3 Desafíos en la protección de microrredes

Proteger las microrredes implica desafíos en la integración de sistemas heterogéneos, gestión de la complejidad y ciberseguridad. La interconexión con la red eléctrica principal expone a vulnerabilidades. Además, la gestión de datos sensibles es crucial. Abordar estos desafíos requiere medidas técnicas y operativas sólidas, incluyendo protocolos de seguridad robustos, monitoreo constante y respuesta rápida a posibles amenazas. La educación y concienciación sobre ciberseguridad son fundamentales. Garantizar la confiabilidad del suministro eléctrico en microrredes es vital para la resiliencia energética y la protección de infraestructuras críticas en un mundo cada vez más interconectado y dependiente de la energía [22].

3.4 Mecanismos de protección y detección

3.4.1 Criptografía y seguridad de la comunicación

La criptografía desempeña un papel crucial en la protección de la comunicación en microrredes. La implementación de algoritmos de cifrado robustos garantiza la confidencialidad e integridad de los datos transmitidos entre los diversos componentes de la microrred. Además, la autenticación basada en criptografía verifica la identidad de los nodos de la red, previniendo la entrada no autorizada [22].

3.4.2 Sistema de detección de intrusiones (IDS)

Los sistemas de detección de intrusiones monitorean continuamente el tráfico de red en busca de patrones sospechosos o comportamientos anómalos. En el contexto de las microrredes, los IDS pueden identificar actividades inusuales, como intentos de acceso no autorizado o patrones de comunicación anómalos, activando alertas y respuestas automáticas para mitigar posibles amenazas [23].

El presente trabajo enfocará sus simulaciones de ciberataques implementando un sistema de detección de intrusiones mediante algoritmos basados en modelo, monitorizando la microrred en un periodo de tiempo definido durante el cual se producirán los hipotéticos ciberataques.

3.4.3 Gestión de identidad y acceso

La gestión de identidad y acceso controla quién tiene acceso a los recursos de la microrred y qué acciones puede realizar. La autenticación multifactor (MFA) y la gestión de privilegios garantizan que solo los usuarios autorizados tengan acceso a sistemas críticos [24]. Esto reduce el riesgo de intrusiones al limitar el acceso a aquellos que realmente necesitan interactuar con la microrred.

La autenticación multifactor es uno de los métodos de gestión de identidad más utilizados actualmente, su funcionamiento consiste básicamente en requerir al usuario alguna verificación adicional de identidad para asegurar que el inicio de sesión o acceso es legítimo (Figura 3.2). Las verificaciones pueden ser variadas, aprovechando que la gran mayoría de usuarios disponen de un dispositivo móvil persona, un sistema de verificación útil es la introducción de un código que es enviado exclusivamente al móvil del usuario de manera que se asegura así la legitimidad del proceso.

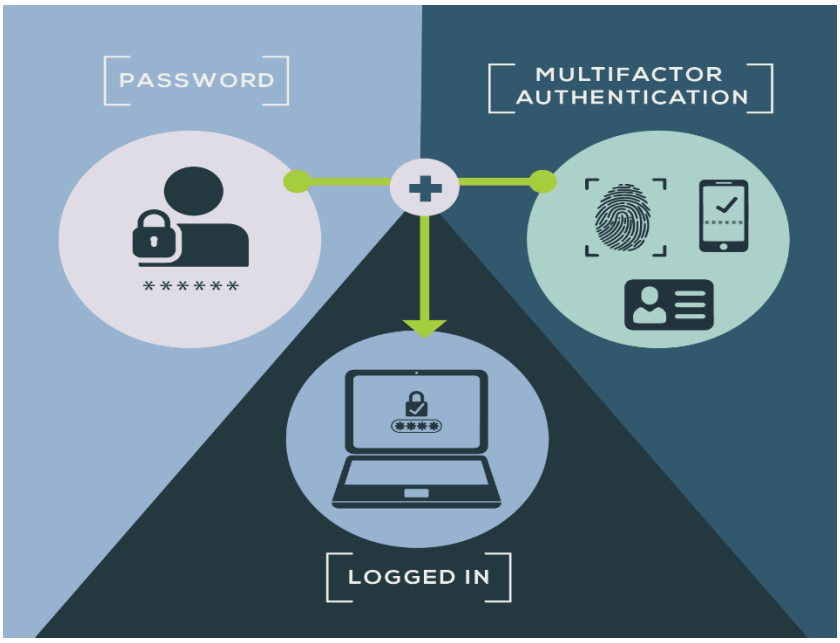

Figura 3.2: *Sistema de autenticación multifactor. Fuente: [24]*

3.4.4 Actualizaciones y parches de seguridad

Es fundamental mantener los sistemas actualizados con los últimos parches de seguridad con el objetivo de mitigar vulnerabilidades conocidas. Las actualizaciones regulares del software y firmware de los componentes de la microrred garantizan que cualquier vulnerabilidad recién descubierta se aborde rápidamente, fortaleciendo la postura de seguridad de la microrred frente a posibles amenazas.

La implementación de estos mecanismos de protección y detección forma un enfoque integral para salvaguardar la seguridad de las microrredes. Combinando tecnologías avanzadas y buenos hábitos de seguridad, se puede llegar a crear una defensa robusta contra las amenazas digitales en un entorno energético cada vez más conectado y digitalizado.

3.5 Estudios de casos reales

3.5.1 Ciberataques a microrredes documentados

Analizar casos específicos de ciberataques a microrredes proporciona una visión práctica de las amenazas y vulnerabilidades que pueden surgir en este entorno específico. Algunos casos han involucrado manipulación remota de sistemas de gestión de energía, comprometiendo la operación autónoma de la microrred. Otros han abordado vulnerabilidades en la capa de comunicación, permitiendo la interceptación de datos críticos [25]. Las empresas energéticas se han vuelto más propensas a los ataques cibernéticos en medio de la pandemia de COVID-19, ya que los atacantes se han

esforzado por beneficiarse de la prisa en la implementación de sistemas remotos y las instalaciones sin personal suficiente. Las compañías eléctricas deben comprender los nuevos riesgos cibernéticos relacionados con el trabajo desde el hogar, como los ataques de ingeniería social y las conexiones a Internet menos confiables, para poder establecer defensas de referencia y limitar las consecuencias de los ciberataques. Los sistemas existentes de las eléctricas están cada vez más conectados a través de sensores y redes y, debido a su naturaleza dispersa, son aún más difíciles de controlar. Esto potencialmente brinda una oportunidad para que los atacantes apunten a la red, similar al ataque en Ucrania ocurrido en diciembre de 2015, donde los piratas informáticos atacaron a tres empresas de distribución de energía en el país, interrumpiendo temporalmente el suministro eléctrico. A medida que las infraestructuras de las eléctricas se vuelven más interconectadas, inteligentes y descentralizadas, un enfoque centralizado para asegurarlas es difícil y será cada vez más insostenible. El monitoreo y la supervisión centralizados son esenciales, pero no suficientes, ya que un sistema central no puede reaccionar con la suficiente rapidez a las amenazas, especialmente cuando el control se fragmenta en numerosos sistemas como las microrredes.

En respuesta a estos ciberataques, las organizaciones y comunidades afectadas han tenido que adaptar rápidamente sus estrategias de seguridad. La implementación de sistemas de detección de intrusiones más avanzados, la mejora de monitorización continua y la colaboración con expertos en ciberseguridad han sido pasos clave para fortalecer la resiliencia de las microrredes. Además, la concienciación y formación del personal se han intensificado para reducir el riesgo de ataques de ingeniería social y phishing.

3.5.2 Lecciones aprendidas y mejores prácticas

El estudio de ciberataques anteriores a microrredes ofrece valiosas lecciones. La identificación temprana de amenazas, la respuesta rápida y la adaptación de estrategias de seguridad son componentes clave de la resiliencia cibernética.

Mediante el estudio de casos reales de ciberataques a microrredes se obtiene una visión más profunda de las amenazas específicas y brinda ideas valiosas sobre cómo mejorar las estrategias de seguridad y fortalecer la seguridad de las microrredes ante futuros ciberataques.

3.6 Herramientas y recursos en ciberseguridad

3.6.1 Prevención y detección

En el campo de la ciberseguridad, la elección de herramientas adecuadas es crucial para la prevención y detección de posibles amenazas. En microrredes, donde la diversidad de componentes y la interconexión son comunes, la implementación de soluciones robustas es esencial. Herramientas como firewalls avanzados, sistemas de detección de intrusiones (IDS) específicos para entornos de energía distribuida y soluciones antivirus adaptadas a sistemas embebidos son fundamentales. Además, el uso de herramientas de análisis de

tráfico y monitorización en tiempo real contribuye a la identificación temprana de actividades sospechosas (Figura 3.3).

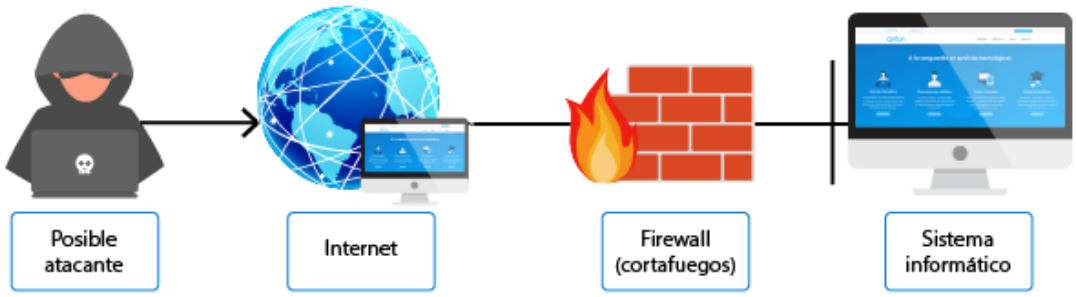

Figura 3.3: *Funcionamiento de un Firewall. Fuente: [26]*

3.6.2 Plataformas de entrenamiento y simulación

La capacitación y la simulación son elementos esenciales en la preparación de enfrentar ciberataques. Plataformas de entrenamiento específicas para ciberseguridad en microrredes permiten a los profesionales adquirir experiencia en entornos simulados. Estas plataformas brindan la oportunidad de practicar la detección y respuesta a amenazas en un entorno controlado, mejorando la preparación del personal ante posibles ciberataques reales. La simulación de incidentes también ayuda a afinar los procesos de respuesta y a evaluar la eficacia de las medidas de seguridad implementadas.

La selección y utilización efectiva de estas herramientas y recursos son esenciales para fortalecer la postura de la ciberseguridad de las microrredes. La combinación de tecnologías avanzadas y entrenamiento continuo permite a las organizaciones anticipar y responder de manera proactiva a las amenazas cibernéticas emergentes en el entorno específico de las microrredes.

Actualmente la formación en ciberseguridad está altamente demandada por gran cantidad de empresas ligadas al mundo tecnológico, lo que provoca que existan multitud de opciones didácticas al alcance de todos sumado a la oportunidad laboral que esto supone [27].

3.7 Desarrollos futuros y tendencias en ciberseguridad

3.7.1 Innovaciones tecnológicas en protección

El futuro de la ciberseguridad en microrredes se perfila con innovaciones tecnológicas destinadas a fortalecer la protección contra amenazas emergentes. El desarrollo de sistemas de inteligencia artificial y aprendizaje automático aplicados a la detección de patrones anómalos permitirá una respuesta más rápida y precisa ante ataques. Además, la implementación de tecnologías de cifrado cuántico podría elevar la seguridad de comunicación a un nuevo nivel que resista incluso ataques más sofisticados de los distintos tipos expuestos en este documento [28].

3.7.2 Colaboración entre entidades

La colaboración de distintas empresas, instituciones académicas y organismos gubernamentales es fundamental para afrontar las amenazas cibernéticas de manera efectiva. La creación de plataformas y mecanismos para compartir información sobre amenazas y vulnerabilidades permitirá una respuesta más coordinada y la aplicación de medidas preventivas a nivel global. La colaboración también facilitará el desarrollo de estándares comunes y marcos regulatorios específicos para la ciberseguridad en microrredes [29].

4 Microrred Experimental del Laboratorio de Energía de El Arenosillo (CEDEA)

El objetivo del presente trabajo está centrado en el estudio de los ciberataques a una microrred, para ello se ha seleccionado la microrred experimental de El Arenosillo, de la cual se dispone de su modelo implementado en Simulink®, y a partir del cual se realizarán las modificaciones que permitirán desarrollar las simulaciones que se correspondan con posibles ciberataques a la microrred.

En este capítulo se hará una descripción general de la microrred a estudiar y se expondrán brevemente cuales son los elementos que la componen.

El CEDEA [1] es el principal campo de pruebas instrumentado para la experimentación de vehículos aeroespaciales del Ministerio de Defensa de España. El CEDEA pertenece al INTA [30] (Instituto de Técnica Aeroespacial) y se sitúa en las proximidades del casco urbano de Mazagón, en el municipio de Moguer (provincia de Huelva, España).

A unos kilómetros del CEDEA se está construyendo el Centro para Ensayos, Entrenamiento y Montaje de Aeronaves no Tripuladas (CEUS). El conjunto CEDEA-CEUS se convertirá en el mejor Centro de Excelencia europeo de Sistemas no Tripulados (Drones), y en referencia internacional para la experimentación con vehículos no tripulados.

Sus principales líneas investigadoras están dirigidas a las energías renovables, la investigación de la atmósfera alta y la colaboración con las Fuerzas Armadas y unidades militares de otros países.

El CEDEA es una instalación flexible con operatividad 24 horas, certificado por la OTAN, la UE y la Agencia Espacial Europea (ESA) como Oficina de Control de Seguridad. El nivel máximo de clasificación para la información y los datos almacenados es OTAN SECRET/ EU SECRET/ ESA SECRET mientras que las áreas de acceso restringido son de clase II. Sus principales funciones, en la actualidad, son:

- Experimentación de cohetes de nuevo desarrollo
- Realización de experimentos científicos con cohetes de sondeo y globos
- Investigaciones atmosféricas
- Pruebas de desarrollo distintos tipos de aeronaves no tripuladas (drones), hasta 150 kg, desde plataforma
- Realización de programas I+D, estudios de durabilidad y ensayos de componentes y sistemas de energía solar

El CEDEA dispone de un laboratorio de energía que cuenta con una microrred experimental, la cual es utilizada para realizar estudios y evaluaciones de distintos campos: control, almacenamiento de energía, sensorización, comunicación y telemonitorización.

4.1 Descripción general de la microrred

La microrred está dispuesta por sistemas de generación eléctrica, sistemas de almacenamiento y cargas en corriente continua y alterna. Existe un bus interno de 408 VDC que conecta los diversos componentes. A su vez, cuenta con una conexión a la red de 230 VAC que proporciona energía al laboratorio. En la Figura 4.1 se muestra un esquema de la distribución y conexiones de los diversos componentes de la microrred en estudio y en la Figura 4.2 se observan los paneles fotovoltaicos, el aerogenerador y el edificio donde se sitúan los componentes principales.

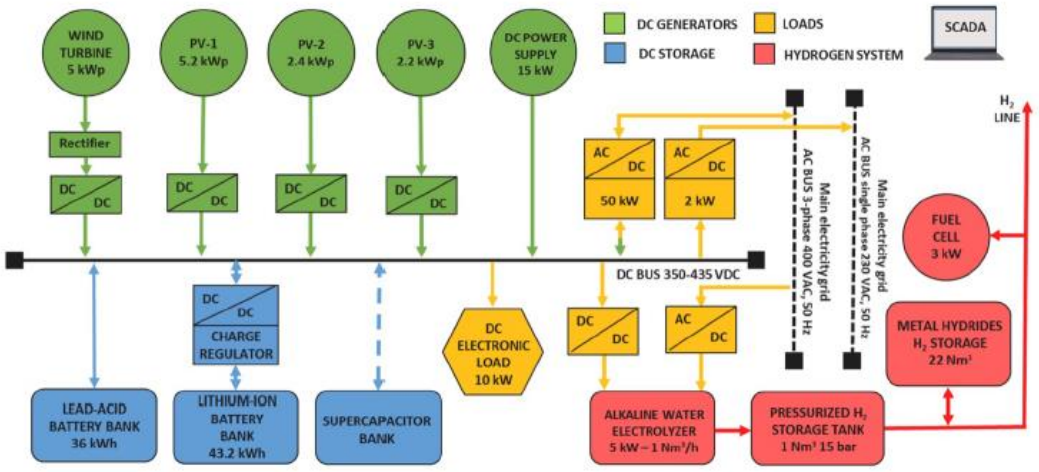

Figura 4.1: *Diagrama conceptual microrred CEDEA. Fuente: [30]*

Figura 4.2: *Vista aérea microrred CEDEA. Fuente: [30]*

4.2 Componentes de la microrred experimental

Para la simulación de posibles ciberataques, los distintos componentes de esta microrred serán los elementos a utilizar para llevar a cabo estas simulaciones, las cuales se detallarán más adelante.

4.2.1 Instalaciones fotovoltaicas

La microrred consta de cuatro campos fotovoltaicos (Figura 4.3):

- **Campo 1:** Está compuesto por 136 paneles BP60 monocristalinos y tiene una potencia máxima total de 5 kWp. Cuenta con cuatro convertidores DC/DC elevadores de tensión o de tipo Boost de 2 kWp cuya salida se conecta al bus general de DC.
- **Campo 2:** Compuesto por paneles dispuestos sobre superficie vertical que están destinados a integración arquitectónica. Existen cinco módulos del modelo ESF-M-BIPVGGP156- 40-161W de la marca Solar Innova. La potencia total es de 2.415 kWp y se conecta mediante convertidores Boost de 3 kWp.
- **Campo 3:** s Este campo está compuesto por paneles flexibles sobre una superficie inclinada. La microrred cuenta con 16 módulos de la marca ENECOM, modelo HF135. En total se obtiene una potencia de 2.16 kWp.Al igual que el campo 2, se conecta con convertidores Boost de 3 kWp.

- **Campo 4:** Este último está formado por 15 módulos policristalinos de paneles inclinados de la marca Wuxi modelo SI-ESF-M- P156-125W. La potencia total es de 1.875 kWp y se conecta directamente a la red mediante inversores que transforman la corriente continua en corriente alterna.

Figura 4.3: *Campos fotovoltaicos de la microrred. Fuente: [30]*

4.2.2 Aerogenerador

Junto a los campos fotovoltaicos, como se puede apreciar en la Figura 4.2, se encuentra el aerogenerador trifásico de eje horizontal de la marca ESHIA modelo AERO5000W de 5 kWp (Figura 4.4). El aerogenerador necesita una etapa de rectificación y elevación para la correcta conexión al bus. De este modo, es posible aumentar la obtención de energía en días nublados. Los diferentes perfiles de generación dependerán del estado del viento, el sol y las nubes.

4.2.3 Fuente de alimentación programable

Además de los sistemas de generación anteriormente mencionados, se dispone de una fuente de alimentación programable de la marca SORENSEN y modelo SGI 500-30D-

1C, capaz de entregar una potencia máxima de 1.5 kW (Figura 4.4). Cuenta con diferentes modos de operación: voltaje constante, corriente constante, potencia constante, protección contra sobrevoltaje, rampa de voltaje y rampa de corriente. La fuente de alimentación se utilizará en los casos de exceso de demanda en los que la generación y el aporte de las baterías no sean suficientes. Como el simulador del que se dispone no presenta ningún riesgo, este componente no se encuentra incluido.

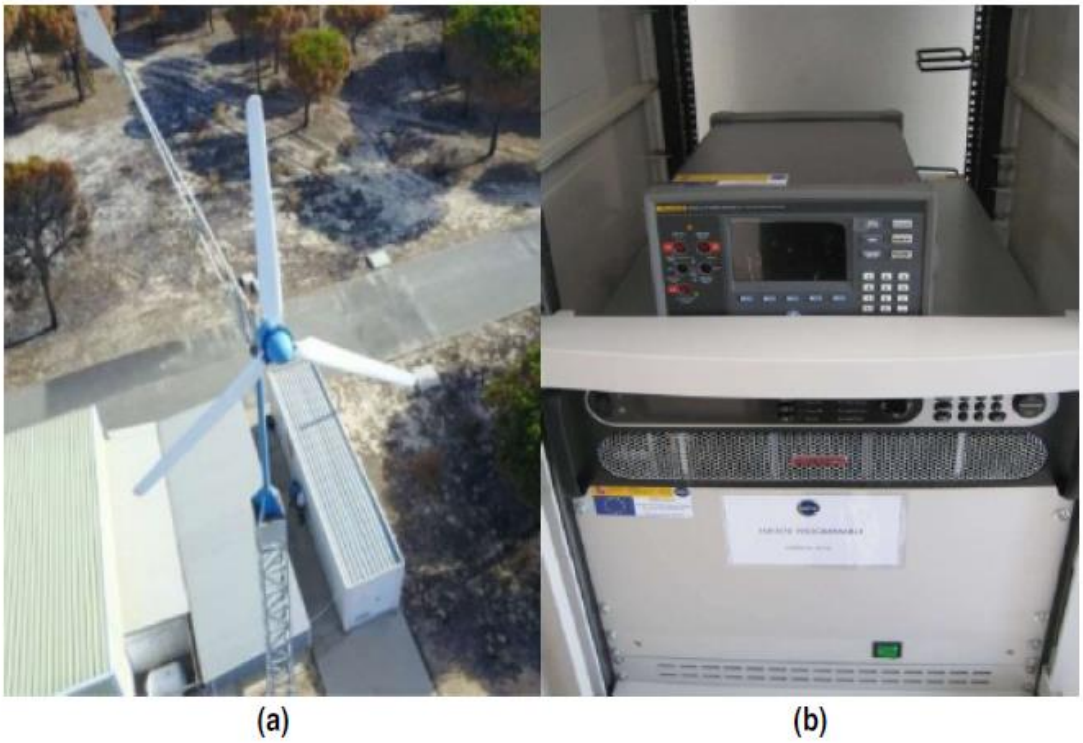

Figura 4.4: *Aerogenerador y fuente de alimentación. Fuente: [30]*

4.2.4 Banco de baterías plomo-ácido

El banco de baterías de plomo-ácido se encuentra directamente conectado al bus de DC y su función es mantener la tensión de operación en él absorbiendo o cediendo la energía necesaria. En la instalación, se pueden encontrar 30 baterías modelo UP-100 del fabricante U-Power con tecnología VRLAAGM (Figura 4.5).

4.2.5 Banco de baterías Ion-Litio

La bancada de baterías de ion litio se conecta al bus de DC a través de un convertidor reductor-elevador bidireccional de 40 kWp. El equipo está formado por 75 celdas de tecnología LiFePO4 modelo ROOK3 del fabricante CEGASA (Figura 4.5), con una capacidad de almacenamiento de 43.2 kWh.

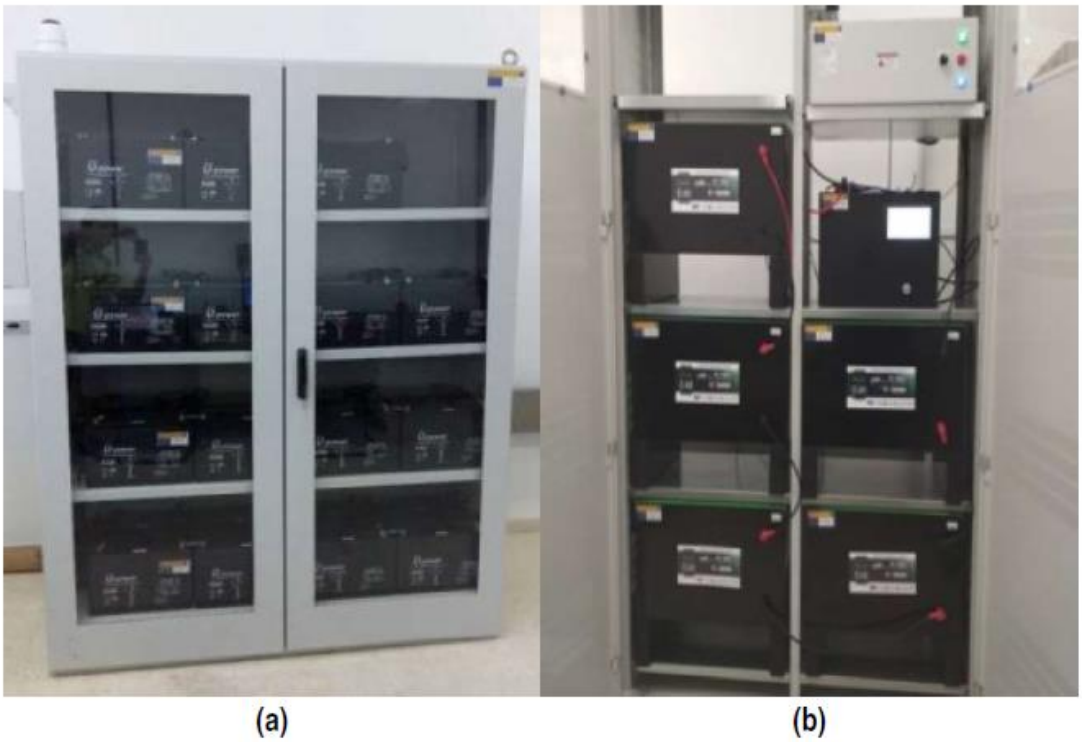

Figura 4.5: *Baterías de plomo y litio. Fuente: [30]*

4.2.6 Electrolizado

La microrred cuenta también con un electrolizador de tecnología alcalina de 30 celdas de la marca ARIEMA (Figura 4.6). Posee una potencia máxima de 5 kW y una ratio máxima de producción de hidrógeno de 1 Nm3/h a 15 bar. Se conecta mediante un convertidor reductor de tensión o Buck y un rectificador, de manera que permite un suministro eléctrico en continua y en alterna.

El hidrógeno producido se almacena en un depósito a baja presión y puede ser utilizado de diferentes maneras: se puede almacenar en hidruros metálicos, utilizar en ensayos de pila de combustible en el laboratorio y comprimir para la recarga de vehículos con pila de combustible.

Figura 4.6: *Electrolizador y depósito de hidrógeno. Fuente: [30]*

4.2.7 Carga electrónica programable

Se dispone de una carga electrónica programable (Figura 4.7) de 10 kW en corriente continua de la marca ADAPTATIVE POWER SYSTEMS modelo 5VP10-32. Consta de dos rangos de potencia: uno inferior de 1 kW y otro superior de 10 kW. Los modos de operación son voltaje constante, corriente constante, potencia constante, resistencia constante, modo combinado de corriente y voltaje constante y modo combinado de voltaje y potencia constante. Además, es posible programar diferentes perfiles de carga dinámicos mediante software del fabricante. Al igual que la fuente programable, la carga podrá ser utilizada para proteger a los equipos ante situaciones en las que la generación y la demanda se encuentren desacompasadas.

4.2.8 Vehículos híbridos

Se dispone de dos vehículos híbridos con baterías y pilas de combustible que pueden ceder o absorber potencia del sistema: Melex y Delfín (Figura 4.7). La plataforma Melex está compuesta por baterías de 48 V y una pila de combustible de 1.2 kWp. Por otro lado, el vehículo Delfín dispone de baterías de 48 V y una pila de combustible de 3 kW. Ambos tienen una capacidad de almacenamiento de hidrógeno de unos 5 Nm3.

Los vehículos se pueden cargar mediante la red de alterna desde el punto de recarga situado en la instalación o pueden inyectar energía a la microrred a través de inversores. De este modo, se implementa el concepto del Vehículo A la Red (Vehicle To Grid (V2G)).

4.2.9 Conexión a la red

Una microrred puede funcionar en modo aislado o en modo no aislado, es decir, puede conectarse y desconectarse de la Red Eléctrica de España (REE) y mantener su funcionamiento. Esta microrred, al igual que la mayor parte de microrredes, se conecta a baja tensión.

Para habilitar el uso de la microrred en modo no aislado, se permite la inyección de potencia a la misma mediante inversores monofásicos limitados a 2 kW. Además, se ha añadido la posibilidad de realizar un intercambio bidireccional con la red mediante un inversor limitado a 50 kW.

Figura 4.7: *Carga electrónica programable y Vehículos híbridos Melex y Delfín. Fuente: [30]*

4.2.10Supercondensador

Se dispone de un banco de supercondensadores, conformado por 7 módulos. Los supercondensadores son de la marca Maxwell y modelo BMOD0141 P064 B04. Proveen de una capacidad de 560kWh, con un voltaje máximo de 64V, corriente máxima de 2000A y una capacidad total de 14.28F.

5 Fundamentos de la detección de fallos y relación con ciberataques

La detección de fallos en el contexto de las microrredes es un componente crítico para garantizar la operación segura y eficiente de estas infraestructuras energéticas descentralizadas.

La monitorización constante de los componentes de la microrred, como generadores distribuidos, sistemas de almacenamiento y dispositivos de gestión, es esencial. Esto implica la recopilación y análisis en tiempo real de datos operativos para identificar cualquier desviación de los parámetros normales de funcionamiento.

Para poder llegar a la conclusión de que se podría estar produciendo un hipotético ciberataque, hay que pasar primero por la detección y aislamiento (localización) de fallos, por ellos el orden lógico de esta secuencia sería el siguiente:

1. **Detección:** llegar a la conclusión de que hay existencia de un fallo (decisión binaria, existe fallo o no existe fallo).
2. **Aislamiento:** determinar que componente de la instalación ha fallado.

El concepto de detección y aislamiento es conocido como FDI [31] (*Fault Detection and Isolation*).

La relación del concepto de fallo con el de ciberataque viene dada porque el ciberataque es o podría ser una sucesión de fallos provocados y revertidos en un periodo de tiempo, siguiendo patrones o no. El "ciberatacante" puede tener multitud de objetivos tal y como se ha explicado anteriormente, ya sean desde robo de datos hasta el simple hecho de denegar el servicio interrumpiendo el correcto funcionamiento de la microrred.

De esta manera se puede llegar a la conclusión de que, para comprender el funcionamiento de un ciberataque en el entorno de una microrred y particularmente el modelo que se va a estudiar en este documento, hay que explicar el concepto de detección de fallo y sus fundamentos.

5.1 Conceptos básicos

A continuación, se definirán una serie de conceptos que se utilizarán a lo largo de este trabajo [32],[33].

- Fallo (*fault)*: Desviación no estándar de al menos una propiedad característica o parámetro del sistema.

- Avería (*failure)*: Interrupción permanente de la capacidad de un sistema para mantener una función requerida bajo condiciones de operación específicas.
- Detección de fallo: Determinación del fallo y el instante de tiempo en el que se ha producido.
- Aislamiento del fallo: Determinación de la localización exacta del fallo.
- Diagnóstico del fallo: Concepto que engloba las etapas de detección, aislamiento e identificación del fallo.
- Modos de fallo: Descripción matemática de los tipos de fallos que puede presentar un componente del sistema.
- Efecto de fallo: Descripción de la propagación del fallo del sistema.
- Sistema de protección: Equipos diseñados para evitar daños personales o materiales a raíz de un fallo.
- Tolerancia de fallos: Capacidad de mantener los objetivos de control a pesar de la aparición de un fallo. Se acepta una degradación relativa en las prestaciones del sistema.
- Supervisor: Entidad que realiza la supervisión de un proceso mediante el diagnóstico de fallos y la determinación de las acciones correctoras que deberán tomarse en presencia de fallos.
- Redundancia física: Exceso de instrumentos para lograr una determinada función.
- Redundancia analítica: Exceso de medios para determinar una variable donde al menos un medio utiliza un modelo matemático.
- Mitigar un fallo: Acción de atenuar o suavizar la ocurrencia de un fallo.
- Acomodación al fallo: Mecanismo de tolerancia a fallos que adapta los parámetros del controlador o de su estructura para evitar los efectos de un fallo. Se alcanzan los objetivos de control, aunque de forma degradada.
- Robustez en la detección de fallos: Capacidad del sistema de diagnóstico de fallos de ser insensible a los errores de modelado, incertidumbres, perturbaciones y ruidos, siendo a su vez sensible a los fallos.
- Robustez activa: Pretende alcanzar la robustez en la generación del residuo.
- Robustez pasiva: Pretende alcanzar la robustez en la toma de decisiones.
- Índice de bondad: Es el intervalo entre el instante que sucede el fallo y el instante en que es detectado.

5.2 Métodos tradicionales de detección de fallos

Enfoques tradicionales para el diagnóstico de fallos anteceden al uso de la redundancia analítica (y, por lo tanto, a la aproximación basada en modelos), aunque hoy en día se siguen utilizando ampliamente. Estos enfoques se basan en técnicas de procesamiento de señales y/o redundancia paralela. Son los siguientes [34]:

- **Análisis del espectro de frecuencia:** Las medidas de las señales de salida tienen un espectro de frecuencia estándar bajo en condiciones de operación estándar, cualquier desviación se puede considerar como una anormalidad. Algunos tipos de fallos pueden incluso tener una anomalía característica en el espectro de manera que pueden ser aislados como fallos solamente con la información

extraída del espectro. Un espectro se puede comparar con un espectro estándar para diagnosticar la naturaleza y tipo de fallo e incluso prever el resultado que el fallo tendrá en el sistema. El análisis espectral es especialmente útil cuando la información del modelo de la planta no está disponible o es difícil de obtener.

- **Enfoque del diccionario de fallos:** En el diccionario de fallos, cada tipo conocido de fallo tiene un comportamiento característico especial en el sistema. Por lo tanto, se puede construir un diccionario de fallos que contenga todos los "comportamientos característicos" conocidos de los fallos. Podemos saber si un fallo ocurre en el sistema al comparar el comportamiento del sistema con repertorios de fallos almacenados en el diccionario de fallos. Este es un método fuera de línea y se puede utilizar para el análisis de fallos después del evento, pero no es adecuado para el diagnóstico de fallos en línea, que es esencial para sistemas críticos de seguridad.
- **Verificación de límites:** Al comparar las variables del proceso con límites predefinidos, donde la superación de un límite puede indicar una situación defectuosa. Aunque es muy simple, este método tiene una seria desventaja en que las variables del proceso pueden variar de acuerdo con diferentes estados operativos del proceso. Por lo tanto, el límite debe depender del estado operativo del proceso. Las ventajas de los métodos basados en la verificación de límites son la simplicidad y la confiabilidad. Sin embargo, solo pueden reaccionar después de un cambio relativamente grande en la característica a evaluar, lo que significa que ha ocurrido un fallo repentino importante o que ha habido una serie de fallos suaves o lentos durante mucho tiempo.
- **Redundancia paralela (hardware):** Un requisito previo importante para lograr la tolerancia a fallos en sistemas de control críticos para la seguridad es mediante líneas múltiples de hardware idéntico. Dos desventajas principales de este enfoque son la penalización de peso (particularmente importante en el control de vuelo) y la posibilidad de que un fallo de modo común del sistema permanezca sin detectar. La redundancia de hardware se puede combinar con métodos basados en modelos para maximizar la eficacia del aislamiento de fallos.

5.3 Principios básicos de la detección de fallos basados en modelo

En las últimas dos décadas, el enfoque de diagnóstico de fallos basado en modelos ha recibido una atención creciente tanto en la investigación como en la. Este enfoque se basa en el concepto de redundancia analítica, en contraposición a la redundancia física (hardware o paralela), que utiliza mediciones de sensores redundantes con fines de diagnóstico de fallos. La redundancia analítica hace uso de señales generadas por el modelo matemático del sistema que se está considerando. Estas señales se comparan con las mediciones reales obtenidas del sistema. La comparación se realiza utilizando cantidades residuales que representan la diferencia entre las señales medidas y las señales generadas por el modelo matemático. Por lo tanto, el diagnóstico de fallos basado en modelos se puede definir como la determinación de fallos en un sistema a partir de la

comparación de las mediciones disponibles del sistema con información a priori representada por el modelo matemático del sistema mediante la generación de cantidades residuales y su análisis. Un residuo es un indicador de fallo o una señal que resalta la situación defectuosa del sistema monitorizado.

La principal ventaja del enfoque de redundancia analítica en comparación con la redundancia de hardware es que no se necesitan componentes de hardware adicionales para implementar un algoritmo de FDI (detección y aislamiento de fallos). Un algoritmo de FDI basado en modelos puede implementarse básicamente en la computadora de control del proceso, a menudo sin requisitos adicionales de hardware. Además, las mediciones necesarias para controlar el proceso son, en muchos casos, también suficientes para el algoritmo de FDI, por lo que no es necesario instalar sensores adicionales. En estas circunstancias, solo se necesita capacidad de almacenamiento adicional y posiblemente una mayor potencia de computación para la implementación de un algoritmo de FDI basado en modelos. El algoritmo de FDI solo requiere procesamiento de datos de entrada y salida para su implementación.

El tema central de los enfoques basados en modelos para FDI es el diseño de señales residuales que llevan información sobre las ubicaciones de fallos y sus momentos de ocurrencia. La figura 5.1 ilustra la estructura conceptual de un sistema de diagnóstico de fallos basado en modelos que comprende dos etapas principales:

- **Generación de residuo:** Las salidas y entradas del sistema son procesadas por un algoritmo apropiado (un procesador) para generar residuos. El residuo debe ser diferente de cero cuando ocurre un fallo y cero en caso contrario. El sistema para generar el residuo se llama generador de residuos.
- **Evaluación de residuos (toma de decisiones):** Los residuos se examinan para determinar la probabilidad de fallos, y luego se aplica una regla de decisión para determinar si ha ocurrido algún fallo.

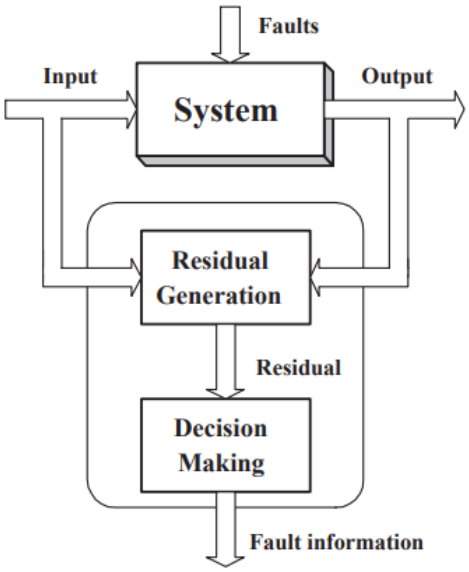

Figura 5.1: *La estructura de dos etapas del FDI basado en modelos. Fuente:[34]*

5.4 Propiedades de los residuos

Los residuos de se componen de las siguientes propiedades fundamentales [31],[33],[35]:

- **Detectabilidad:** Capacidad de detectar un fallo en el proceso monitorizado. Dado que en la práctica ocurren errores de modelado, perturbaciones y ruidos no es fácil diferenciar e identificar la ocurrencia real de fallos de estas posibles perturbaciones, por lo que es necesario recurrir al uso de umbrales. Es destacable que la utilización de umbrales muy amplios puede llegar a enmascarar fallos, por lo que la determinación de la magnitud correcta de estos umbrales debe ser analizada. Lo ideal sería diseñar sistemas de diagnóstico robustos que eviten en la medida de lo posible las interferencias debidas a los ruidos, perturbaciones o errores de modelado.
- **Capacidad de aislamiento:** La exitosa detección de un fallo es seguida por el procedimiento de aislamiento del fallo, el cual distinguirá (aislará) un fallo específico de los demás. Mientras que un solo residuo es suficiente para detectar fallos, se requiere un conjunto de residuos (o un residuo en forma de vector) para el aislamiento del fallo. Si un fallo es distinguible de otros fallos mediante uno o más residuos, se puede definir como un fallo aislable (Figura 5.2). Para facilitar el aislamiento, generalmente se genera un conjunto de residuos (o una forma vectorial de residuo, de una de las siguientes maneras:

    - **Conjunto de residuos estructurados:** En este método, se genera un conjunto de residuos (figura 5.3). Cada residuo está diseñado para ser sensible a diferentes fallos o subconjuntos de fallos, mientras que es insensible a los fallos restantes.

- **Vector residual de dirección fija:** Otra forma de mejorar la capacidad de aislamiento es diseñar un vector residual direccional que se ubique en una dirección fija y específica de fallo (o subespacio) en el espacio residual en respuesta a un fallo particular (figura 5.4).

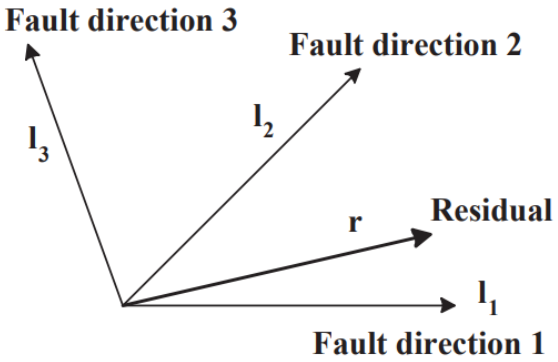

Figura 5.2: *El concepto del vector residual direccional para el aislamiento de fallo. Fuente: [34]*

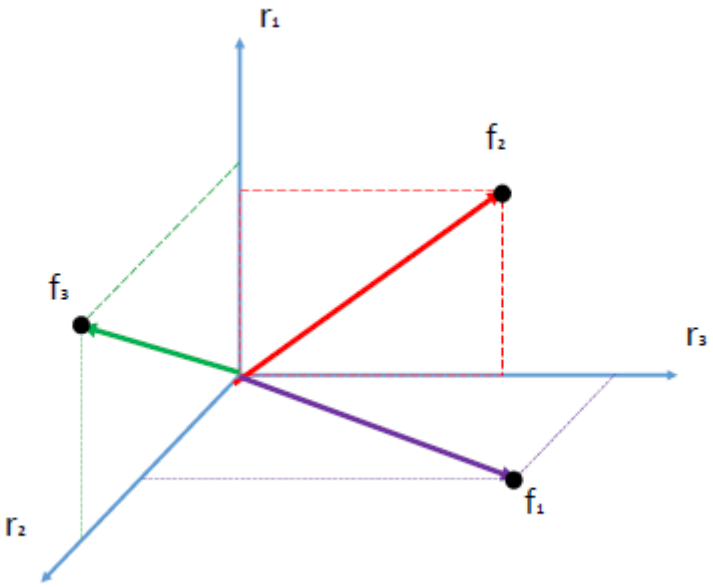

Figura 5.3: *Ejemplo de residuo estructurado. Fuente: [35]*

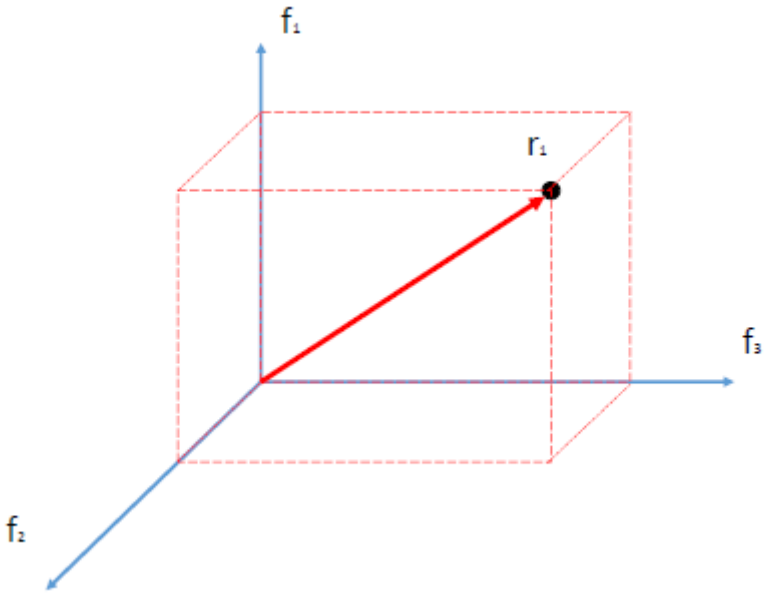

Figura 5.4: *Ejemplo de residuo dirigido. Fuente: [35]*

5.5 Métodos de Generación de Residuos Basados en Modelos

En esta sección se procederá a definir y explicar los diferentes métodos que permiten llevar a cabo una generación de residuos que permita detectar y aislar fallos.

5.5.1 Enfoques basados en observadores

El principio básico de los enfoques basados en observadores es estimar la salida del sistema en función de las entradas y salidas del sistema monitoreado a través de un observador, y el residuo es la diferencia (ponderada) entre las salidas estimadas y reales.

Las propiedades fundamentales de este método son [31]:

- La tarea del aislamiento puede llevarse a cabo mediante un conjunto de residuos estructurados o un vector residual direccional.
- La reacción ante fallos es muy rápida
- Muy adecuado para detectar y aislar fallos en actuadores y sensores
- El proceso de diseño es sistemático y sencillo
- Fácil de implementar y el algoritmo de ejecución es simple
- Fácil manejo de múltiples fallos si el número de medidas es suficiente
- Métodos para manejar el ruido del sistema:
  - Propiedades estadísticas desconocidas: Se puede aplicar un filtro adicional al residuo, basado en suposiciones sobre las bandas de frecuencia de fallo y ruido.

    - Propiedades estadísticas conocidas: Se puede utilizar un filtro de Kalman para producir el residuo libre de fallos con mínima varianza y, por lo tanto, reducir las alarmas falsas y perdidas
- Robustez frente a perturbaciones desconocidas
- Robustez frente a errores de modelado

5.5.2 Enfoque de relaciones de paridad

El principio básico de los enfoques de relaciones de paridad para FDI es verificar la inconsistencia entre las entradas y salidas del sistema monitorizado. El término de desequilibrio se utiliza como una señal residual. Se ha demostrado que el enfoque de relaciones de paridad es equivalente al uso de un observador de respuesta nula [31]. Una señal residual generada por un observador no dé respuesta nula es equivalente a una señal residual post filtrada que es generada por un observador de respuesta nula. Debido a la correspondencia entre enfoques basados en observadores y enfoques de relaciones de paridad, la mayoría de sus condiciones de aplicabilidad son las mismas.

5.5.3 Enfoques de estimación de parámetros

El principio básico consiste en estimar los parámetros del modelo en línea y el residuo se basa luego en una comparación entre los valores estimados y reales de los parámetros. Alternativamente, el error entre la salida real y la predicha (generada por la estimación de parámetros en tiempo real) puede utilizarse como una señal residual.

Las características principales de este método de detección de fallos son las siguientes [36]:

- La reacción ante fallos es lenta.
- La detección y aislamiento de fallos en actuadores y sensores son posibles pero complicados.
- La detección y aislamiento de fallos en parámetros son muy directos.
- El procedimiento de diseño es sistemático, pero no simple.
- Complejidad de implementación (requiere gran cantidad de cálculos).
- La detección y aislamiento de múltiples fallos no es una tarea fácil a menos que se instale un gran número de sensores.
- El ruido es fácil de manejar en el procedimiento de estimación de parámetros.
- No linealidad, posible de manejar mediante técnicas de identificación para sistemas no lineales.
- Robustez, dependiente de las propiedades de robustez del método de estimación.
- Excelente capacidad adaptativa y de autoaprendizaje.

5.5.4 Enfoque de redes neuronales

Los enfoques basados en redes neuronales para la generación de residuos son similares a los enfoques basados en observadores, aunque el observador se reemplaza por una red neuronal. La salida del sistema se estima a través de la red neuronal, y el residuo se obtiene a partir de la diferencia entre las salidas estimadas y reales. Las características importantes de los enfoques basados en redes neuronales se pueden resumir de la siguiente manera [36]:

- La reacción ante fallos incipientes es moderada.
- Muy conveniente para detectar y aislar fallos en actuadores y sensores.
- Buena capacidad para detectar y aislar fallos en parámetros.
- Es posible la detección de múltiples fallos.
- Tolerancia al ruido moderada.
- Requieren extensas sesiones de entrenamiento durante el diseño.
- Complejidad de implementación, requiere una gran cantidad de cálculos
- Excelente para manejar sistemas no lineales, que es para lo que están diseñadas las redes neuronales.
- No hay investigaciones publicadas sobre problemas de robustez aún.
- Sin requisitos a priori para el modelado.
- Capacidad adaptativa y de autoaprendizaje excelente

5.6 Relación entre ciberataques y detección de fallos

La relación entre ciberataques y la detección de fallos en una microrred puede entenderse a través de la vulnerabilidad de los sistemas de control y la importancia de contar con métodos efectivos de detección de fallos para garantizar la seguridad y la integridad del sistema. Existe un artículo [37] que proporciona una perspectiva sobre cómo el control predictivo basado en modelo puede ser utilizado para detectar y mitigar los efectos de los ciberataques en sistemas de control, lo que puede ser relevante para explorar la relación entre la detección de fallos y la ciberseguridad. Aquí hay algunas consideraciones:

1. **Interrupción del sistema de control:**

    - **Ciberataques:** Un ataque cibernético podría comprometer los sistemas de control de una microrred, causando interrupciones en las operaciones normales.
    - **Detección de fallos:** Los métodos de detección de fallos pueden identificar anomalías en el comportamiento del sistema que podrían ser indicativas de un ataque.

2. **Manipulación de datos de sensores:**

    - **Ciberataques:** Un atacante podría manipular los datos de los sensores para engañar al sistema de control y causar malfuncionamientos.
    - **Detección de fallos:** La detección de fallos puede identificar discrepancias entre los datos medidos y los valores esperados, indicando posibles manipulaciones.
3. **Inyección de comandos maliciosos:**

    - **Ciberataques:** Inyectar comandos maliciosos en el sistema de control puede causar operaciones no deseadas.
    - **Detección de fallos:** La detección de fallos puede identificar cambios abruptos en el comportamiento del sistema que podrían ser causados por comandos maliciosos.
4. **Robo de información confidencial:**

    - **Ciberataques:** Los ataques pueden estar dirigidos a robar información confidencial sobre la configuración y el funcionamiento de la microrred.
    - **Detección de fallos:** La detección de fallos puede ayudar a identificar intentos no autorizados de acceder a información confidencial.
5. **Degradación de rendimiento:**

    - **Ciberataques:** Algunos ciberataques pueden estar diseñados para degradar el rendimiento de la microrred.
    - **Detección de fallos:** La detección de fallos puede alertar sobre cambios en el rendimiento que podrían indicar un ataque.

6. **Anomalías de comunicación:**

    - **Ciberataques:** Ataques a la infraestructura de comunicación pueden afectar la capacidad del sistema para recibir información crítica.
    - **Detección de fallos:** La detección de fallos puede identificar interrupciones en la comunicación y alertar sobre posibles problemas.

La combinación de ciberseguridad y detección de fallos es esencial para garantizar la confiabilidad y la seguridad de las microrredes, especialmente en entornos críticos como sistemas de energía distribuida.

6 Métodos de detección y aislamiento de fallos utilizados en la microrred

6.1 Método de ecuaciones de paridad

El concepto básico del diagnóstico de fallos basado en modelos consiste en generar una estimación de las salidas a partir del modelo de la planta, de forma que pueda ser evaluada la consistencia del sistema modelado y el sistema real en cada instante de tiempo. Cuando ocurre un fallo, se produce una inconsistencia entre el sistema modelado y el sistema real. Esta inconsistencia, llamada residuo, se puede expresar como [35]:

$$r(t) = y(t) - \hat{y}(t) \tag{6.1}$$

Siendo $r(t)$ el residuo, $y(t)$ la salida real del sistema e $\hat{y}(t)$ la salida estimada.

El modelo lineal en espacio de estados para un sistema discreto ideal sin perturbaciones, ruidos ni fallos es:

$$x(t+1) \;\; = \; Ax(t) + Bu(t) \tag{6.2}$$

$$y(t) \;\;\; = \;\;\; Cx(t) + Du(t) \tag{6.3}$$

Siendo $x(t)$ el vector de estados, $u(t)$ el vector de entradas del proceso e $y(t)$ el vector de salidas del sistema. A, B, C, y D son las matrices que representan el comportamiento de la planta en el punto de operación, $A \in \mathbb{R}^{nxn}$, $B \in \mathbb{R}^{nxp}$, $C \in \mathbb{R}^{qxn}$, $D \in \mathbb{R}^{qxp}$, siendo:

- n = número de variables
- p = número de entradas
- q = número de salidas

Esta formulación puede simplificarse considerando $D = 0$

$$x(t+1) \;\; = \; Ax(t) + Bu(t) \tag{6.4}$$

$$y(t) = Cx(t) \tag{6.5}$$

Introduciendo la ecuación (6.4) en (6.5) para $(t+1)$ se obtiene:

$$y(t+1) = CAx(t) + CBu(t) \tag{6.6}$$

Repitiendo el proceso para $(t+2)$ resulta:

$$y(t+2) = CA^2x(t) + CABu(t) + CBu(t+1) \tag{6.7}$$

Y de forma general para $(t+p1)$:

$$y(t+p1) = C(A^{p1}x(t) + \sum_{i=0}^{p1-1}(A^{n1-1-i}Bu(t+i))) \tag{6.8}$$

De esta forma se obtiene la salida estimada para una ventana de tiempo $(t+p1)$, siendo $p1 \leq n$. Se puede expresar en su forma compacta como:

$$Y(t+p1) = Ox(t) + T_uU(t+p1) \tag{6.9}$$

Y para un número de muestras desplazadas en el tiempo hacia atrás por $p1$:

$$Y(t) = Ox(t-p1) + T_uU(t) \tag{6.10}$$

En la ecuación (6.10) aparecen las entradas y salidas, así como el vector de estado inicial para $(t+p1)$, generando una redundancia temporal, donde:

$$O = [C \quad CA \quad CA^2 \quad \cdots \quad CA^{p1}]^T \tag{6.11}$$

$$T_u = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{p1-1}B & CA^{p1-2}B & \cdots & 0 \end{bmatrix} \tag{6.12}$$

$$Y(t) = [y(t-p1) \quad y(t-p1+1) \;\; y(t-p1+2) \;\; \cdots \;\; y(t)]^T \tag{6.13}$$

$$U(t) = [u(t-p1) \quad u(t-p1+1) \;\; u(t-p1+2) \;\; \cdots \;\; u(t)]^T \tag{6.14}$$

La matriz O es la matriz de observabilidad y la matriz T_u es la matriz de Toeplitz.

Reordenando la ecuación (6.14) se tiene:

$$Y(t) - T_u U(t) = Ox(t-p1) \tag{6.15}$$

Llamando al primer término forma computacional:

$$Y(t) - T_u U(t) \tag{6.16}$$

Y al segundo, forma interna:

$$Ox(t-p1) \tag{6.17}$$

Volviendo a la ecuación (6.1) y siendo la ecuación (6.10) la salida estimada, el residuo se puede calcular como:

$$r(t) = Y(t) - [Ox(t-p1) + T_u U(t)] \tag{6.18}$$

Se puede observar en la ecuación (6.18) que el residuo depende del estado $x(t)$. Lo ideal es que no existiera esta dependencia, por lo que se multiplica la ecuación (6.18) por un vector w tal que:

$$wO = 0 \tag{6.19}$$

Por tanto, si esta condición se satisface, el residuo queda desacoplado del estado y queda la expresión del residuo como:

$$r(t) = wY(t) - wT_u U(t) \tag{6.20}$$

Para obtener más residuos se aumenta el número de vectores w formando una matriz W, quedando el vector de residuos:

$$r(t) = WY(t) - WT_u U(t) \tag{6.21}$$

El orden de W determina el número de ecuaciones de paridad.

Si se supone la existencia de perturbaciones, ruidos y fallos en el sistema, las ecuaciones que representan el modelo lineal en espacio de estado serían:

$$x(t+1) = Ax(t) + Bu(t) + Ev(t) + Lf(t) \tag{6.22}$$

$$y(t) = Cx(t) + Du(t) + Hv(t) + Mf(t) \tag{6.23}$$

Donde las matrices $E, L, H\ y\ M$ se incluyen para modelar el comportamiento del sistema en el punto de operación cuando se considera la existencia de perturbaciones, ruidos y fallos. $f(t)$ es el vector de fallos aditivos y $v(t)$ el vector de perturbaciones y ruidos.

Operando de la misma manera que en (5.8) se obtiene:

$$\begin{aligned} y(t+p1) = C(A^{p1}x(t) + \sum_{i=0}^{p1-1}\left(A^{n1-1-i}Bu(t+i)\right) + \sum_{i=0}^{p1-1}\left(A^{n1-1-i}Ev(t+i)\right) \\ + \sum_{i=0}^{p1-1}\left(A^{n1-1-i}Lf(t+i)\right) + Hv(t+p1) + Mf(t+p1) \end{aligned} \tag{6.24}$$

En su forma compacta para un tiempo $p1$ desplazado hacia atrás queda:

$$Y(t) = Ox(t-p1) + T_u U(t) + T_v V(t) + T_f f(t) \tag{6.25}$$

En este caso las matrices de Toeplitz $T_v\ y\ T_f$ para una ventana de tiempo $p1$ quedaría:

$$T_v = \begin{bmatrix} H & 0 & \cdots & 0 \\ CE & H & \cdots & 0 \\ CAE & CE & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{p1-1}E & CA^{p1-2}E & \cdots & H \end{bmatrix} \quad T_f = \begin{bmatrix} M & 0 & \cdots & 0 \\ CF & M & \cdots & 0 \\ CAF & CF & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{p1-1}F & CA^{p1-2}F & \cdots & M \end{bmatrix} \tag{6.26}$$

Y los vectores $V(t)$ y $F(t)$

$$V(t) = [v(t-p1)\ v(t-p1+1)\ \ v(t-p1+2)\ \cdots v(t)]^T \tag{6.27}$$

$$F(t) = [f(t-p1)\ f(t-p1+1)\ \ f(t-p1+2)\ \cdots f(t)]^T \tag{6.28}$$

Nuevamente, para que el residuo no depende del estado se ha de cumplir la expresión (6.19). Una vez satisfecha, el cálculo del residuo sería:

$$r(t) = wY(t) - [wT_u U(t) + wT_v V(t) + wT_f f(t)] \tag{6.29}$$

Y para una matriz W :

$$r(t) = WY(t) - [WT_u U(t) + WT_v V(t) + WT_f f(t)] \tag{6.30}$$

6.2 Estimación de estado mediante observadores

Un modelo lineal en espacio de estados de un proceso multivariable viene descrito por:

$$x(t+1) = Ax(t) + Bu(t) \tag{6.31}$$

$$y(t) = Cx(t) \tag{6.32}$$

Suponiendo que se tiene una entrada $u(t)$ de dimensión p y una salida $y(t)$ de dimensión q, el observador se utiliza para reconstruir las variables de estado que no son medidas a partir de las entradas y salidas al modelo.

La ecuación del observador es

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + H[y(t) - C\hat{x}(t)] \tag{6.33}$$

Donde $\hat{x}(t+1)$ es el estado estimado para $t+1$, $\hat{x}(t)$ es el estado estimado para t, $u(t)$ la entrada manipulable, $y(t)$ la salida del sistema y H es la matriz de ganancias del observador.

El error en la salida se define como:

$$e(t) = y(t) - C\hat{x}(t) \tag{6.34}$$

Reordenando los términos en (6.33) se obtiene:

$$\hat{x}(t+1) = [A - HC]\hat{x}(t) + Bu(t) + Hy(t) \tag{6.35}$$

Donde se asume que el sistema es observable, es decir:

$$O = [C \;\; CA \;\; CA^2 \;\; \cdots \;\; CA^{n-1}] \tag{6.36}$$

O tiene rango máximo, siendo C_{qxn} y A_{nxn}

El error en el estado se describe como la diferencia entre el estado real y el estado estimado (6.37)

$$\tilde{x}(t+1) = x(t+1) - \hat{x}(t+1) \tag{6.37}$$

Introduciendo en (6.37) las ecuaciones (6.32) y (6.35) se obtiene:

$$\tilde{x}(t+1) = [A - HC]\hat{x}(t) \tag{6.38}$$

Consecuentemente, el error en el estado converge asintóticamente

$$\lim_{t\to\infty} \tilde{x}(t) = 0 \tag{6.39}$$

Para cualquier desviación de estado $= [x(0) - \tilde{x}(0)]$ si el observador es estable. Esto puede ser conseguido eligiendo de manera adecuada la matriz H, mediante, por ejemplo, posicionamiento de polos.

6.3 Estimación de estado mediante Filtro de Kalman

Como se ha descrito en (6.33), para un sistema representado en espacio de estados dado por (6.32), la ecuación del observador de estado viene dada por:

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + H[y(t) - C\hat{x}(t)] \tag{6.40}$$

A partir de la cual se ha definido una ecuación de error en la salida (6.34) y un error de estimación del estado (6.38).

Si se tiene un proceso sin perturbaciones, el observador converge al estado verdadero del sistema si los autovalores de $[A - HC]$ son asintóticamente estables. En la velocidad de dicha convergencia tienen un papel muy importante los autovalores de la matriz de ganancias del observador H.

Sin embargo, bajo la influencia de perturbaciones estocásticas, la reconstrucción de un estado mediante observadores no es la óptima.

Si al proceso se le suma ruido estocástico $v(t) +$ a la entrada y $n(t)$ a la salida, la ecuación (6.32) quedaría de la siguiente forma:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Vv(t) \\ y(t) &= Cx(t) + n(t) \end{aligned} \tag{6.41}$$

Las matrices del proceso, $A, B, C\ y\ V$ son conocidas. El estado inicial $x(0)$en principio no es conocido, pero se tiene información estadística del mismo y también de $v(t)$ y $n(t)$.

Estas variables se suponen estadísticamente independientes y con una distribución normal gaussiana con los valores medios:

$$E\{x(0)\} = x_0 \quad E\{v(t)\} = 0 \quad E\{n(t)\} = 0 \tag{6.42}$$

Y las matrices de covarianza:

$$E\{(x(0)-x_0)(x(0)-x_0)^T\} = X_0\ E\{v(t)v^T(t)\} = M\ E\{n(t)n^T(t)\} = N \tag{6.43}$$

Donde M y N también se suponen conocidas.

Dado que el error de la estimación no puede converger a cero, se debe llevar a cabo una optimización sobre la estimación del vector de estados $x(t)$, basándonos en las entradas $u(t)$e $y(t)$. Hay que llevar a cabo una optimización por mínimos cuadrados:

$$\min\left|\left|x(t)-\hat{x}(t|j)\right|\right|^2 \tag{6.44}$$

En (6.44) se utilizan dos marcos temporales. t hace referencia al tiempo actual y j al instante de tiempo de las medidas. La estimación del estado recibe varios nombres:

- k>j: Problema de predicción
- k=j: Problema de filtrado
- k<j: Problema de estabilidad

En los problemas de predicción y de filtrado, las medidas de las salidas utilizadas es la siguiente:

$$Y_j = \{y(0), y(1), y(2), \dots y(j)\} \tag{6.45}$$

Se utiliza la siguiente notación:

- Estimaciones óptimas:

$$\hat{x}(t|j) = E\{x(t)|Y_j\} \tag{6.46}$$

- Error de la estimación:

$$\hat{x}(t|j) = x(k) - \hat{x}(t|j) \tag{6.47}$$

- Matrices de covarianza del error de la estimación

$$P^{-}(t+1) = E\{\tilde{x}(t+1|t)\tilde{x}^{T}(t+1|t)\} \qquad (6.48)$$
$$P(t+1) = E\{\tilde{x}(t+1|t+1)\tilde{x}^{T}(t+1|t+1)\}$$

Para el instante $t+1$ la variable de estado $x(t+1)$ puede ser predicha utilizando el modelo en espacio de estados (6.41) con la información en el instante t.

$$\hat{x}(t+1|t) = A\hat{x}(t|t) + Bu(t) + V\bar{v}(t) \qquad (6.49)$$

Se utiliza $\underline{v}(t)$ ya que v(t) no es conocida.

Suponiendo que $E\{v(t)\} = \underline{v} = 0$ quedaría:

$$\hat{x}(t+1|t) = A\hat{x}(t|t) + Bu(t) \qquad (6.50)$$

En el instante $t+1$ también está disponible la salida $y(t+1)$, siendo:

$$y(t+1) = Cx(t+1) + n(t+1) \qquad (6.51)$$

Sin embargo, $x(t+1)$ es desconocida. La predicción $\hat{x}(t)$ está perturbada por el ruido $v(t)$ y la salida medible $y(t+1)$ por $n(t+1)$.

Si tanto $\hat{x}(t+1|t)$ como $x(t+1)$ fuesen conocidos, se podría calcular el estado como una media ponderada:

$$\hat{x}(t+1|t+1) = (I_n - K')\hat{x}(t+1|t) + K'x(t+1) \qquad (6.52)$$

Reordenando términos en (6.52):

$$\hat{x}(t+1|t+1) = \hat{x}(t+1|t) + K'[x(t+1) - \hat{x}(t+1|t)] \qquad (6.53)$$

Donde K' es una matriz de ponderación de orden n (número de estados) que se elige para minimizar la covarianza del error de estimación $P(t-1)$. Ahora, haciendo $K' = KC$ convertimos el vector de estados $x(t+1)$ en el vector de salidas medibles $y(t+1)$. Aplicando esta transformación en (6.52):

$$\hat{x}(t+1|t) = [I_n - KC]\hat{x}(t+1|t+1) + Ky(t+1) \tag{6.54}$$

La ecuación (6.54) contiene:

- $\hat{x}(t+1|t)$: La predicción del modelo de $x(t+1)$ basado en la última estimación $\hat{x}(t|t)$
- $y(t+1)$: La nueva medida

A (6.52) le sigue un algoritmo de estimación recursivo:

$$\hat{x}(t+1|t+1) = \hat{x}(t+1|t) + K(t+1)[y(t+1) - C\hat{x}(t+1)|t)] \tag{6.55}$$

Donde la matriz de corrección $K(t+1)$ ha debido ser elegida para minimizar la matriz de covarianza del error de estimación. Como esta covarianza varía en el tiempo, $K(t+1)$ también debe hacerlo.

El error en la estimación es:

$$\tilde{x}(t+1|t) = \hat{x}(t+1|t) - E\{x(t+1)\} \tag{6.56}$$

Y el error en la medida es:

$$\tilde{y}(t+1) = \hat{y}(t+1) - E\{y(t+1)\} = n(t) \tag{6.57}$$

Las correspondientes matrices de covarianzas son:

$$\begin{aligned} P^{-}(t+1) &= E\{\tilde{x}(t+1|t)\tilde{x}^T(t+1|t)\} \\ Y &= E\{\tilde{y}(t+1)\tilde{y}^T(t+1)\} \end{aligned} \tag{6.58}$$

La matriz de covarianza de la estimación recursiva $\tilde{x}(t+1)$ viene por tanto dada por:

$$P(t+1) = [I_n - K(t+1)C]P^{-}(t+1)[I_n - K(t+1)C]^T + K(t+1)NK^T(t+1) \tag{6.59}$$

Ahora se busca un valor de $K(t+1)$ que minimice la varianza de la covarianza del error de estimación:

$$K(t+1) = P^{-}(t+1)C^{T}[CP^{-}(t+1)C^{T} + N]^{-1} \tag{6.60}$$

Junto con:

$$P(t+1) = P^{-}(t+1) - K(t+1)CP^{-}(t+1) \tag{6.61}$$

Donde P^{-}, que es la matriz de covarianzas del error de estimación $\tilde{x}(t|t)$, se obtiene de:

$$P^{-}(t+1) = AP(t)A^{T} + VMV^{T} \tag{6.62}$$

A modo resumen, la secuencia de cálculos es:

1. **Predicción:** de (6.49) y (6.52)

$$\begin{aligned} \hat{x}(t+1|t) &= A\hat{x}(t|t) + Bu(t) \\ P^{-}(t+1) &= AP(t)A^{T} + VMV^{T} \end{aligned} \tag{6.63}$$

2. **Corrección:** de (6.60), (6.55) y (6.61)

$$\begin{aligned} K(t+1) &= P^{-}(t+1)C^{T}[CP^{-}(t+1)C^{T} + N]^{-1} \\ \hat{x}(t+1|t+1) &= \hat{x}(t+1|t) + K(t+1)[y(t+1) - C\hat{x}(t+1)|t)] \\ P(t+1) &= [I_n - K(t+1)C]P^{-}(t+1) \end{aligned} \tag{6.64}$$

Y por último, introduciendo la predicción en la corrección queda:

$$\hat{x}(t+1) = Ax(t) + Bu(t) + K(t+1)[y(t+1) - C\big(A\hat{x}(t) + Bu(t)\big)] \tag{6.65}$$

La estimación del estado es una estimación recursiva del estado $\hat{x}(t+1|t+1)$ basado en el estado predicho $\hat{x}(t+1|t)$ por el modelo del proceso y una corrección basada en la nueva medida $y(t+1)$.

La matriz de ganancias K depende de las matrices de covarianzas M y N.

Si las matrices del sistema A, B, C y las matrices de covarianza de los ruidos no dependen del tiempo, la ganancia del Filtro de Kalman, $K(t+1)$ converge asintóticamente a un valor estable. La matriz de covarianza del estado estable del error de estimación P^{-} viene dado por:

$$P^{-}(t+1) = AP^{-}(t) - AP^{-}(t)C^{T}[CP^{-}(t)C^{T} + N]^{-1}CP^{-}(t)A^{T} + VMV^{T} \tag{6.66}$$

La cual es una ecuación de Riccati. Su solución asintótica proporciona el valor estable de P^-. Por tanto, el valor estable de la ganancia de Kalman:

$$\underline{K} = P^- C^T [C P^- C^T + N]^{-1} \tag{6.67}$$

La secuencia de cálculos se reduciría a:

1. **Predicción:**

$$\hat{x}(t+1|t) = A\hat{x}(t|t) + Bu(t) \tag{6.68}$$

2. **Corrección:**

$$\hat{x}(t+1|t+1) = \hat{x}(t+1|t) + \overline{K}[y(t+1) - C\hat{x}(t+1)|t)] \tag{6.69}$$

Introduciendo la predicción en la corrección queda:

$$\hat{x}(t+1|t) = A\hat{x}(t|t-1) + Bu(t) + A\overline{K}[y(t) - C\big(A\hat{x}(t|t-1)\big)] \tag{6.70}$$

6.4 Definición de los Umbrales Estocásticos

En un algoritmo de generación de residuos ideal, el valor de éste debería ser nulo en ausencia de fallos y no nulo cuando se produjese alguno. Sin embargo, debido a los diversos motivos que se han mencionado anteriormente, el valor de residuo no es por lo general nulo incluso en ausencia de fallos.

Es entonces cuando surge la necesidad de diseñar unos márgenes o umbrales para determinar si el sistema está funcionando correctamente o puede existir la presencia de algún fallo.

Utilizando la Teoría de la Probabilidad, se definen los Umbrales Estocásticos [35], teniendo en cuenta la incertidumbre en la generación de residuos para optimizar los valores de umbrales de fallos. Por otro lado, la definición de estos umbrales trae consigo el inconveniente de que puedan existir pequeños fallos que no sean detectados, al no superar los umbrales definidos, lo que lleva a la presencia de posibles falsos negativos.

La Teoría de la Probabilidad se ocupa del estudio de procesos estocásticos en el ámbito matemático. En este contexto, se establece la variable aleatoria como una función matemática que asigna un número a un experimento aleatorio, es decir, busca representar los resultados de un evento aleatorio. En particular, una variable aleatoria continua abarca un conjunto infinito de posibles valores. Por lo tanto, los residuos se tratarán como variables aleatorias continuas. La expresión que describe la probabilidad relativa de que

una variable aleatoria adquiera un valor específico recibe el nombre de Función de Densidad de Probabilidad (Probability Density Function, PDF). Cumple las siguientes condiciones:

1. Para todos los valores de x:

$$f(x) \geq 0 \tag{6.71}$$

2. El área bajo la curva de $f(x)$

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty \leq X \leq \infty) = 1 \tag{6.72}$$

Siendo X el valor de la variable aleatoria, x el valor de la variable real (eje de abscisas) y P la probabilidad que va de 0 a 1.

La *Cumulative Distribution Function (CDF)* es la probabilidad de que una variable aleatoria sea menor o igual a un valor determinado, es decir, siendo X una variable aleatoria continua con una PDF $f(x)$, la CDF de X es:

$$F(x) = \int_{-\infty}^{X} P(X \leq x) \tag{6.73}$$

Para obtener el valor de x asociado con una probabilidad acumulada específica p , se utiliza la *Inverse Cumulative Distribution Function (ICDF).* Para una distribución continua y estrictamente monótona la ICDF devuelve un valor x tal que:

$$P(X \leq x) = p \tag{6.74}$$

Centrándonos en el ámbito de este documento, los residuos se ven afectados por variables aleatorias, lo que generalmente impide que sean nulos, convirtiéndolos en un proceso estocástico. Utilizando los principios de la Teoría de la Probabilidad, sería posible determinar un umbral $\gamma(t)$ con una probabilidad p. Esto indicaría que el residuo generado $r(t)$ tiene una probabilidad igual o mayor a dicho umbral en situaciones donde no hay fallos.

De manera análoga a (6.74):

$$P\big(r(t) \leq \beta(t)\big) = p \tag{6.75}$$

La probabilidad de que ocurra un falso positivo es de $1-p$. De este modo si p es elevada disminuye la probabilidad de que aparezcan falsos positivos, pero aumenta la probabilidad de que se produzcan falsos negativos. Por tanto, es de vital importancia encontrar un equilibrio entre la sensibilidad y especificidad.

Lo que se intenta es calcular para cada instante t el umbral superior, $\beta_i(t)$ e inferior $\gamma_i(t)$ cuyo intervalo debe contener al residuo $r_i(t)$ en ausencia de fallos. Se expresa como:

$$P\big(\gamma_i(t) \le r_i(t) \le \beta_i(t)\big) = F_{r(t)}\big(\beta(t)\big) - F_{r(t)}(\gamma(t)) \tag{6.76}$$

Siendo $F_{r(t)}$ en este caso el valor de la CDF de una distribución normal.

Para calcular los umbrales se utiliza un conjunto de datos históricos experimentales en escenarios ausentes de fallos a partir de los cuales se realiza el cálculo de los residuos históricos.

Matemáticamente el problema se podría resolver como:

$$\begin{gathered} \text{máx} \quad \beta_i(t) \\ \text{s.t. } P(r_i^h(t) \ge \beta_i(t)) \ge 1-p, \forall i, \forall t \end{gathered} \tag{6.77}$$

$$\begin{gathered} \text{mín} \quad \gamma_i(t) \\ \text{s.t. } P(r_i^h(t) \le \gamma_i(t)) \ge p, \forall i, \forall t \end{gathered} \tag{6.78}$$

Donde P es la distribución de probabilidad, $r_i^h(t)$ es el histórico del residuo i, β_i representa el umbral superior del residuo r_i, γ_i el umbral inferior y p representa la probabilidad. Las expresiones (6.77) y (6.78) se pueden escribir como:

$$P(r_i^h(t) \ge \beta_i(t) \ge 1 - F(\beta_i(t)) \rightarrow \beta_i(t) \le F^{-1}(1-p) \tag{6.79}$$

$$P\left(r_i^h(t) \le \gamma_i(t)\right) \ge 1 - F(\gamma_i(t)) \rightarrow \ \gamma_i(t) \ge F^{-1}(p) \tag{6.80}$$

La amplitud del intervalo $[\beta_i(t), \gamma_i(t)]$ determina el tamaño de fallo mínimo que puede ser detectado en valores absolutos.

Para determinar si un residuo $r_i(t)$ distinto de cero revela un fallo o no, una nueva variable binaria $r_i^b \in [0,1]$ se empareja con r_i. Se calcula de acuerdo con:

$$r_i^b(t) = \begin{cases} 1 & si\ r_i(t) < \beta_i(t) \quad o \quad r_i(t) < \gamma_i(t) \\ 0 & si\ \beta_i(t) \le r_i(t) \le \gamma_i(t) \end{cases} \tag{6.81}$$

7 Diseño de una microrred para su implementación en MatLab® y Simulink®

A continuación, se desarrollará el modelo lineal en espacio de estados de la microrred del INTA, el cual será base para los algoritmos de cálculos de residuos implementados, para ello se adaptará el modelo de microrred proporcionado [5],[38].

7.1 Modelo de la microrred

El primer paso para la implementación de un sistema de detección de fallos basado en modelo es obtener un modelo de la microrred que se va a estudiar. Dado que la relación entre la potencia en los diferentes componentes y los estados de carga se asemeja a una función lineal, el sistema se puede modelar mediante una representación en espacio de estados:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Dd(t) \\ y(t) &= Cx(t) \end{aligned} \qquad (7.\ 1)$$

Donde $x(t)$ es el vector de estado, $u(t)$ es la señal de control, $d(t)$ es la perturbación en el sistema e $y(t)$ es la salida.

El estado del sistema viene representado por los estados de carga de las baterías, el supercondensador y el nivel del depósito de hidrógeno, como se observa en (7.2). En este modelo el estado se corresponde con la salida.

$$x(t) = y(t) = [SOC_{pb(t)}\ SOC_{li(t)}\ LOH(t)\ SOC_{sc}(t)\ SOC_{me(t)}\ SOC_{de}(t)]^T \qquad (7.\ 2)$$

Los términos relativos a los estados de carga de los coches híbridos serán nulos mientras no estén conectados a la microrred.

Las baterías de plomo se encargan de permitir que se cumpla el balance de potencia. Por ello, las variables manipulables son las potencias de las baterías, del supercondensador y de la red:

$$u(t) = [P_{grid(t)} \;\; P_{li}(t) \;\; P_{sc}(t) \;\; P_{me}(t) \;\; P_{de}(t)]^T \quad (7.3)$$

Al igual que en (7.2), las componentes relativas a los vehículos híbridos serán distintas de cero cuando estos se encuentren conectados.

Debido a que la generación y la demanda no forman parte del sistema de control, se tratarán como una perturbación del sistema. Además, si los coches híbridos están conectados a la microrred, el consumo de hidrógeno, el cual está expresado en porcentaje, también se verá representado como una perturbación.

$$d(t) = [P_{net}(t) \quad Cons_H(t)]^T \quad (7.4)$$

A continuación, se generará un modelo lineal que representa los diversos componentes que constituyen la microrred, estableciendo una conexión entre los estados y las señales de las variables que pueden ser controladas a través de su rendimiento.

Las baterías de ácido plomo se descargan al aumentar la potencia que intercambian con la microrred y se pueden modelar mediante el siguiente balance de energía:

$$SOC_{pb}(t+1) = SOC_{pb}(t) - \frac{\eta_{pb} T_s}{C_{max_{pb}}} P_{pb}(t) \quad (7.5)$$

Donde $C_{max_{pb}}$ es la capacidad de la batería, η_{pb} es el rendimiento y T_s es el tiempo de muestreo.

Al no ser una señal de control la potencia de la batería de plomo se ha de expresar en función del resto de potencias mediante un balance energético.

$$P_{pb}(t) = -P_{net}(t) - P_{grid}(t) - P_{li}(t) - P_{me}(t) - P_{de}(t) - P_{sc}(t) \quad (7.6)$$

Uniendo las ecuaciones (7.5) y (7.6), se obtiene el modelo de batería de plomo que viene dado por la siguiente ecuación:

$$SOC_{pb}(t+1) = SOC_{pb}(t) - \frac{\eta_{pb} T_s}{C_{max_{pb}}} (-P_{net}(t) - P_{grid}(t) \quad (7.7)$$

La batería de ion litio se modela de manera similar:

$$SOC_{li}(t+1) = SOC_{li}(t) - \frac{\eta_{li} T_s}{C_{max_{li}}} P_{li}(t) \quad (7.8)$$

El estado de carga del supercondensador queda:

$$SOC_{sc}(t+1) = SOC_{sc}(t) - \frac{\eta_{sc} T_s}{C_{max_{sc}}} P_{sc}(t) \quad (7.\ 9)$$

En el modelo del electrolizador, el nivel de hidrógeno se modela como:

$$LOH(t+1) = LOH(t) + \frac{\eta_{elz} T_s}{V_{max}} P_{elz}(t) - Cons_H(t) \quad (7.\ 10)$$

Donde $Cons_H$ es la disminución del nivel de hidrógeno provocada por los coches híbridos.

Por último, las baterías de los coches se modelan del mismo modo que las de litio y plomo:

$$SOC_{me}(t+1) = SOC_{me}(t) - \frac{\eta_{me} T_s}{C_{max_{me}}} P_{me}(t) \quad (7.\ 11)$$

$$SOC_{de}(t+1) = SOC_{de}(t) - \frac{\eta_{de} T_s}{C_{max_{de}}} P_{de}(t) \quad (7.\ 12)$$

En la siguiente tabla se recogen los valores de los diferentes parámetros del modelo, donde los valores de los rendimientos se han obtenido de forma experimental y el resto son datos obtenidos del catálogo:

Tabla 7.1 *Parámetros del modelo*

	η	$C_{max}(Ah)$
Batería plomo	0.2000 (6% V^{-1})	125
Batería litio	0.2318 (6% V^{-1})	225
Supercondensador	0.0238 (m^3/skW)	125
Plataforma Melex	0.3978 (m^3/skW)	225
Plataforma Delfín	0.3978 (m^3/skW)	225

A partir de lo desarrollado anteriormente, se puede crear el modelo de espacio de estados de la siguiente forma:

$$\begin{bmatrix} SOC_{pb}(t+1) \\ SOC_{li}(t+1) \\ LOH(t+1) \\ SOC_{sc}(t+1) \\ SOC_{me}(t+1) \\ SOC_{de}(t+1) \end{bmatrix} = A \cdot \begin{bmatrix} SOC_{pb}(t) \\ SOC_{li}(t) \\ LOH(t) \\ SOC_{sc}(t) \\ SOC_{me}(t) \\ SOC_{de}(t) \end{bmatrix} + B \cdot \begin{bmatrix} P_{grid}(t) \\ P_{li}(t) \\ P_{sc}(t) \\ P_{me}(t) \\ P_{de}(t) \end{bmatrix} + D \begin{bmatrix} P_{net}(t) \\ Cons_H(t) \end{bmatrix} \quad (7.\ 13)$$

Donde las matrices son las siguientes

$$A = I_6 \quad (7.\ 14)$$

$$B = \begin{bmatrix} \frac{\eta_{pb} T_s}{C_{max_{pb}}} & \frac{\eta_{pb} T_s}{C_{max_{pb}}} & \frac{\eta_{pb} T_s}{C_{max_{pb}}} & \frac{\eta_{pb} T_s}{C_{max_{pb}}} & \frac{\eta_{pb} T_s}{C_{max_{pb}}} \\ 0 & \frac{-\eta_{li} T_s}{C_{max_{li}}} & 0 & 0 & 0 \\ & & \frac{-\eta_{sc} T_s}{C_{max_{sc}}} & 0 & 0 \\ 0 & 0 & & & \\ 0 & 0 & 0 & \frac{-\eta_{me} T_s}{C_{max_{me}}} & 0 \\ 0 & 0 & & & \\ & & 0 & 0 & \frac{-\eta_{de} T_s}{C_{max_{de}}} \end{bmatrix} \quad (7.\ 15)$$

$$C = I_6 \quad (7.\ 16)$$

$$D = \begin{bmatrix} \frac{\eta_{pb} T_s}{C_{max_{pb}}} & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \end{bmatrix} \quad (7.\ 17)$$

7.2 Entradas para los bloques de cálculo de los residuos

El método de cálculo de los residuos es un algoritmo desarrollado en un fichero "script" de MatLab® e implementado en Simulink® mediante un bloque que ejecuta dicha función en MatLab®.

Las entradas a dichos algoritmos son:

1. **Salidas de la microrred:**
   - Estado de carga de la batería de plomo (SOC_{pb}).
   - Estado de carga de la batería de litio (SOC_{li}).

- Estado de carga del supercondensador (SOC_{sc}).
- Nivel de depósito de hidrógeno (LOH).
- Estado de carga de la batería del coche Melex (SOC_{me}).
- Estado de carga de la batería del coche Delfín (SOC_{de}).

2. **Entradas a la microrred:**
   - Potencia de la red (P_{grid}).
   - Potencia de la batería de litio (P_{li}).
   - Potencia del supercondensador (P_{sc}).
   - Potencia de la batería del coche Melex (P_{me}).
   - Potencia de la batería del coche Delfín (P_{de}).
3. **Perturbaciones:**
   - Potencia neta (P_{net}).
   - Consumo de hidrógeno por parte de los vehículos híbridos ($ConsH$).

Con estos elementos y el modelo de la microrred representado en (7.1) podremos obtener las salidas de la microrred en cada muestra de tiempo y obtener un valor residual para cada una de ellas.

7.3 Ecuaciones de paridad

7.3.1 Definición de las matrices del algoritmo

Las matrices utilizadas para calcular los residuos mediante las ecuaciones de paridad son las propias matrices del espacio de estados, con el que se modela el comportamiento de la microrred.

Para proceder de manera adecuada con la programación en MatLab® se debe ser consciente que éste ha de funcionar simultáneamente con la microrred, es decir, ha de proporcionar las entradas e ir proporcionando un valor residual a la vez que las entradas se van generando. El tiempo de procesamiento no puede ser superior al tiempo que transcurre entre cada muestra para que los residuos proporcionados sean válidos, ya que en caso contrario no se corresponderían con el estado actual de la microrred.

Este método se caracteriza por funcionar teniendo en cuenta una ventana de tiempo determinada $p1 \leq n$. Para programar el algoritmo se ha tomado la ventana de tiempo $p1 = n$, siendo n el número de estados igual a 6 (7.2). De este modo las matrices definidas de forma genérica en (6.1) quedan particularizadas para nuestro caso de la siguiente manera:

$$O = [C \quad CA \quad CA^2 \quad \cdots \quad CA^6]^T_{42x6} \tag{7. 18}$$

$$T_u = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^5B & CA^4B & \cdots & 0 \end{bmatrix}_{42x35} \quad T_v = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CE & 0 & \cdots & 0 \\ CAE & CE & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^5E & CA^4E & \cdots & 0 \end{bmatrix}_{42x14} \tag{7. 19}$$

La matriz W, conformada por vectores w, responsable de eliminar la dependencia de las ecuaciones de paridad con el estado del sistema y definida como:

$$wO = 0 \tag{7. 20}$$

Se calcula como:

$$W_{36x42} = null(O^T)^T \tag{7. 21}$$

Es decir, W es el espacio vectorial nulo traspuesto de la matriz de observabilidad traspuesta.

Como se ha descrito anteriormente, este método se desarrolla teniendo en cuenta una ventana de tiempo $p1$. Cuando se inicia el funcionamiento, el algoritmo no es capaz de generar residuos dado que no tiene suficientes muestras, por lo que el valor del residuo en los primeros instantes de tiempo, hasta llegar a $p1$, se toma como nulo.

7.3.2 Definición de las entradas al algoritmo

Como se ha expuesto en la sección anterior, en el comienzo del funcionamiento no se disponen de suficientes muestras como para hacer funcionar el algoritmo de la manera deseada. De este modo, durante los primeros $p1$ instantes de tiempo, la función únicamente se dedica a almacenar de forma estructurada, tal y como requiere el método de las ecuaciones de paridad, un histórico de entradas, salidas y perturbaciones del modelo de la microrred. Una vez transcurrido el tiempo necesario, el algoritmo comienza a generar los valores residuales, eliminando las muestras que ya no necesita y añadiendo las nuevas muestras que alimentan al bloque de la función.

La estructura de las muestras para hacer funcionar el algoritmo se define en (6.13), (6.14) y (6.27) de manera general. Particularizando para el caso en estudio, las variables quedan:

$$Y(t) = \begin{bmatrix} SOC_{pb}(t-6) \\ \vdots \\ SOC_{pb}(t) \\ SOC_{li}(t-6) \\ \vdots \\ SOC_{li}(t) \\ LOH(t-6) \\ \vdots \\ LOH(t) \\ SOC_{sc}(t-6) \\ \vdots \\ SOC_{sc}(t) \\ SOC_{me}(t-6) \\ \vdots \\ SOC_{me}(t) \\ SOC_{de}(t-6) \\ \vdots \\ SOC_{de}(t) \end{bmatrix}_{42x1} \quad U(t) = \begin{bmatrix} P_{grid}(t-6) \\ \vdots \\ P_{grid}(t) \\ P_{li}(t-6) \\ \vdots \\ P_{li}(t) \\ P_{sc}(t-6) \\ \vdots \\ P_{sc}(t) \\ P_{me}(t-6) \\ \vdots \\ P_{me}(t) \\ P_{de}(t-6) \\ \vdots \\ P_{de}(t) \end{bmatrix}_{35x1} \quad V(t) = \begin{bmatrix} P_{net}(t-6) \\ \vdots \\ P_{net}(t) \\ ConsH(t-6) \\ \vdots \\ ConsH(t) \end{bmatrix}_{14x1} \qquad (7.\ 22)$$

A modo de resumen, la expresión que engloba el conjunto de matrices con sus correspondientes dimensiones para obtener el residuo por el método de las ecuaciones de paridad queda de la siguiente manera:

$$Residuo_{6x1} = W_{6x42}Y_{42x1} - [W_{6x42}Tu_{42x35}U_{35x1} + W_{6x42}Tv_{42x35}V_{35x1}] \qquad (7.\ 23)$$

Mediante la expresión (7.23) se obtiene un vector de seis componentes en la que cada componente es el valor residual a cada elemento que conforma la microrred.

7.4 Filtro de Kalman

7.4.1 Adecuación de las matrices para el algoritmo

Para hacer funcionar el observador basado en el filtro de Kalman teniendo en cuenta las perturbaciones del modelo (P_{net} y $ConsH$) hay que reagrupar las matrices del modelo lineal en espacio de estados a diferencia del caso de las ecuaciones de paridad.

En la definición del modelo de la microrred del INTA, se ha obtenido un modelo lineal en el que los estados depende de las señales de control y de las perturbaciones. Para simplificar el modelo y adecuarlo al algoritmo, las perturbaciones se introducirán en el vector de estados. Las nuevas matrices que se obtendrán se llaman matrices ampliadas.

Originalmente, el sistema tiene la forma:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Dd(t) \\ y(t) &= Cx(t) \end{aligned} \qquad (7.\ 24)$$

Pero el algoritmo del filtro de Kalman se ejecuta para un modelo de la siguiente forma:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{7. 25}$$

Por tanto, se van a definir unas nuevas matrices $A, B\ y\ C$, haciendo que el modelo lineal en espacio de estados quede con la estructura siguiente:

$$\begin{bmatrix} x(t+1) \\ d(t+1) \end{bmatrix} = \begin{bmatrix} A & D \\ 0 & I \end{bmatrix} \begin{bmatrix} x(t) \\ d(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) \tag{7. 26}$$

$$y(t) = [C \quad 0] \begin{bmatrix} x(t) \\ d(t) \end{bmatrix} \tag{7. 27}$$

De este modo, ya se obtienen las matrices mostradas en (7.25) teniendo en cuenta las perturbaciones como parte del vector de estados del sistema.

7.4.2 Cálculo de la matriz de ganancias del observador basado en filtro de Kalman

Puesto que las matrices del sistema $A, B\ y\ C$ y las matrices de covarianzas no dependen del tiempo, la ganancia del filtro de Kalman converge asintóticamente a un valor estable.

Para obtener el valor de la matriz de ganancias se debe resolver una ecuación de Riccati (6.66). Para ello, MatLab® proporciona un comando (*dlqr*) que resuelve la ecuación de Riccati y proporciona el valor estable de la matriz de ganancias, que se usará en el término de corrección del filtro de Kalman:

$$\begin{aligned} \hat{x}_{8x1}(t+1|t) = {} & A_{8x8}\hat{x}_{8x1}(t|t-1) + B_{8x5}u_{5x1}(t) \\ & + A_{8x8}\overline{K}_{8x6}\left[y_{6x1}(t) - C_{6x8}\left(A_{8x8}\hat{x}_{8x1}(t|t-1)\right)\right] \end{aligned} \tag{7. 28}$$

Donde $A, B\ y\ C$ son las matrices obtenidas en (7.26) y (7.27).

El valor residual obtenido con el observador basado en el filtro de Kalman se obtiene de:

$$Residuo_{6x1} = y_{6x1} - C_{6x8}\hat{x}_{8x1} \tag{7. 29}$$

Siendo y la salida de la microrred y $\hat{x}$ la estimación del estado mediante el filtro de Kalman.

8 Simulación y resultados

En este capítulo se procederá a realizar las pruebas sobre el modelo de microrred proporcionado para el desarrollo de este trabajo. A continuación, se adjunta la representación del modelo proporcionado por [38] en dos imágenes distintas para su mejor lectura (Figura 8.1 y 8.2):

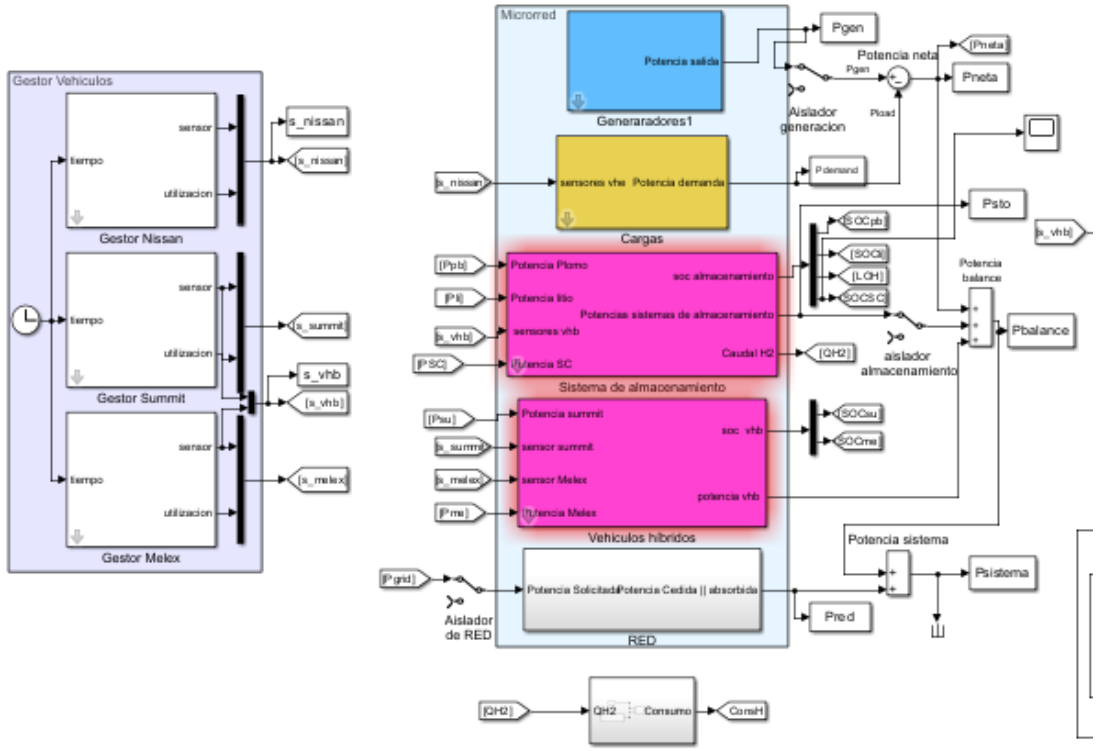

Figura 8.1: *Primera mitad del diagrama del modelo proporcionado incluyendo los algoritmos FDI.*

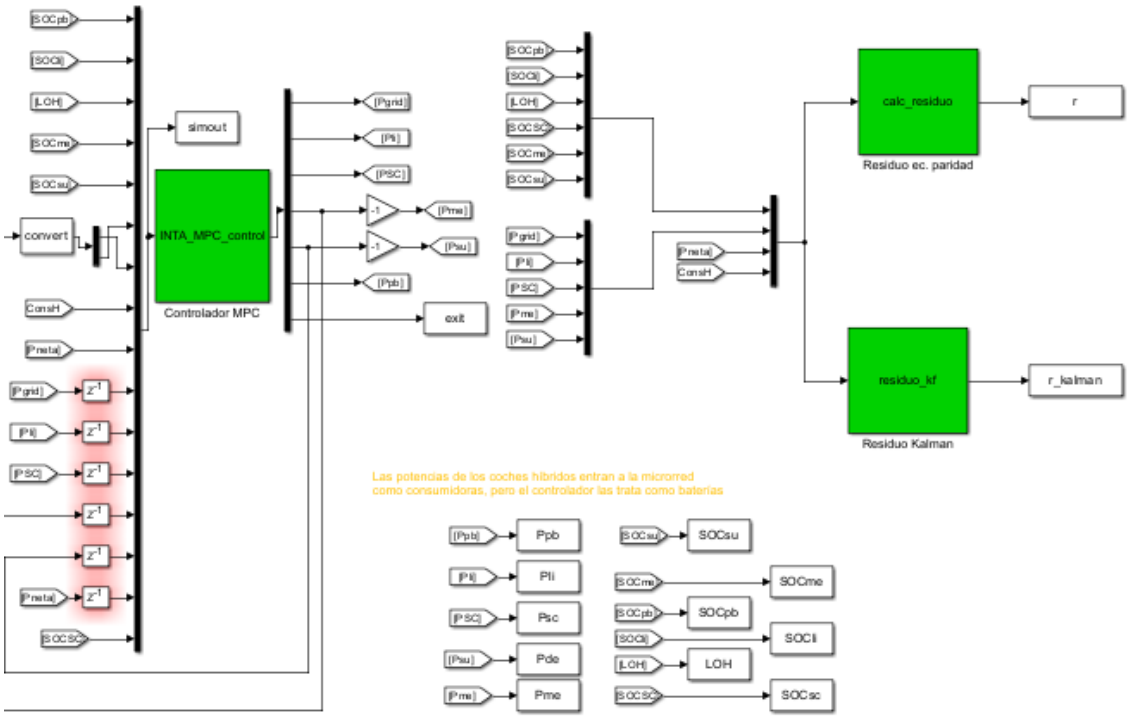

Figura 8.2: *Segunda mitad del diagrama del modelo proporcionado incluyendo el controlador MPC y los dos bloques de detección de fallos mediante la generación de residuos.*

El modelo de microrred representado en las Figuras 8.1 y 8.2 es controlado por un controlador MPC, el cual utiliza las salidas y entradas anteriores de la microrred para generar, optimizando una función de coste, nuevas entradas para el modelo. El controlador proporcionado tiene un tiempo de muestreo de un segundo, un horizonte de control de dos segundos y un horizonte de predicción de diez segundos.

El modelo proporcionado, parte de un modelo previo al cual se le añadieron los dos bloques de detección de fallos mediante la generación de residuos, los cuales pueden apreciarse en la figura (8.2).

El presente trabajo parte del modelo descrito anteriormente, el cual estaba diseñado para la detección de fallos y el posterior aislamiento de manera que se determine en qué momento ocurre el fallo y qué componente es el que está fallando.

Como continuación al modelo de microrred proporcionado, la propuesta de este trabajo trata sobre la detección de posibles ciberataques que pudieran ocurrir, en este caso a los distintos componentes que integran la microrred. La distinción entre un ciberataque y un fallo puede considerarse de la siguiente manera: La ocurrencia de un fallo es algo que sucede de manera puntual, de modo que la alteración del funcionamiento de un

componente en un intervalo de tiempo está marcada por la existencia o no de un cambio en el comportamiento. Sin embargo, cuando se produce un ciberataque, tal y como se ha desarrollado anteriormente en el capítulo 3, el objetivo de los "ciberatacantes" puede ser la extracción de información comprobando los distintos funcionamientos del sistema realizando numerosas acciones en un periodo de tiempo determinado.

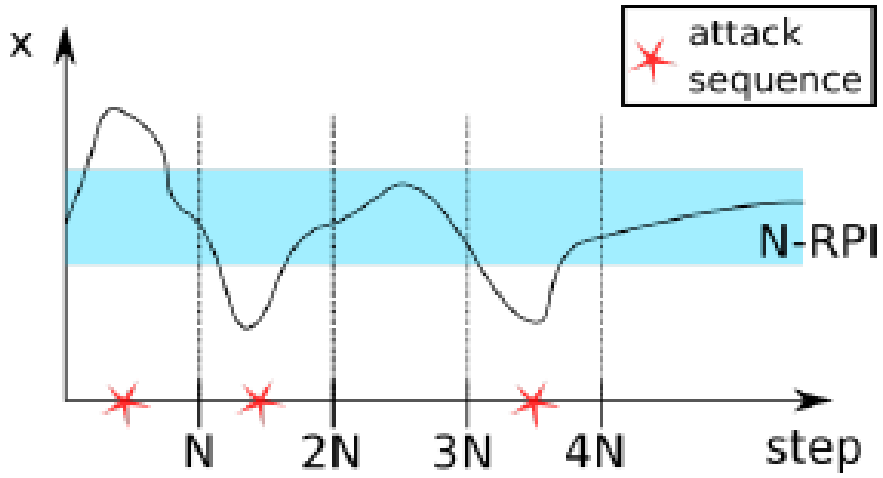

Figura 8.3: *Secuencia de ciberataque a lo largo de N pasos. Fuente: [39]*

La figura 8.3, extraída de un artículo de la Universidad de Burdeos (Université Bourdeaux) que trata sobre ciberataques "DoS" (denegación de servicio) a sistemas controlados por modelos basados en control predictivo MPC, proporciona una visión de cómo sería la secuencia de un ciberataque a lo largo de N pasos. La franja azul se correspondería con lo que sería el "umbral" de nuestro modelo predictivo y la línea de trazo continuo representada en la gráfica se correspondería con el residuo generado. Las estrellas rojas indican la presencia de ciberataque en el paso correspondiente. Como puede observarse, si el modelo está correctamente programado y no hay existencia de falsos positivos ni falsos negativos, cuando hay presencia de ciberataque la gráfica correspondiente al residuo debería de sobrepasar el umbral establecido indicando así la presencia de ciberataques.

La diferencia fundamental entre una secuencia de fallos consecutivos no provocada de manera intencionada con otra provocada intencionadamente es que, una vez ocurrido el fallo, la probabilidad de que el funcionamiento del sistema volviese a ser el adecuado sería mínima, pues necesitaría de la intervención de un factor humano a menos que estuviese programado para su reajuste automático. En un ciberataque, el atacante manipula a su voluntad el funcionamiento de la red, provocando esta sucesión de intervalos en los que se alterna un funcionamiento correcto, dentro del valor de los umbrales establecidos, y un incorrecto funcionamiento que sobrepasa estos umbrales, para finalmente terminar en un correcto funcionamiento, lo que da lugar a sospecha de que el sistema está siendo atacado. Además, los delincuentes informáticos por lo general quieren pasar desapercibidos durante las interferencias que realizan, salvo que el objetivo principal sea desestabilizar el sistema.

Es entonces la base fundamental para la realización de las simulaciones el procedimiento anteriormente explicado, el cual servirá para justificar el modelo de ciberataque que será ejecutado en las simulaciones que se representarán a continuación.

Al igual que en el procedimiento de detección de fallos, en donde se consideraba la existencia de fallo cuando el valor de residuo superaba el valor de un umbral preestablecido, en el caso de ciberataque el funcionamiento será similar, con la particularidad de que un ciberataque podría considerarse una sucesión consecutiva de fallos que ocurren en un periodo determinado de tiempo.

El modelo de simulación proporcionado permite elegir entre el estado de carga inicial, desde carga baja hasta carga media y alta. De cara a los resultados, seleccionando cualquier estado de carga la respuesta de los algoritmos a los fallos es equivalente, de modo que la elección no es relevante a la hora de evaluar los resultados. Se selecciona por tanto un estado de carga medio.

En las secciones siguientes, se evaluará el valor del residuo ante ciberataques provocados en elementos de la microrred teniendo en cuenta la influencia del valor p en la determinación de los umbrales.

Para llevar a cabo la simulación de los ciberataques en el modelo de la microrred, se ha procedido a actualizar el “switch” original del modelo proporcionado, el cual permitía de manera manual provocar un fallo sobre un componente en un determinado instante de tiempo, al conseguir que la potencia facilitada sea nula, lo que indicaría presencia de fallo.

La actualización del “switch” consiste en incluir un bloque función de MatLab® en el que se ha programado una secuencia de fallos y “no fallos”, mediante una secuencia binaria que permite alternar de “no fallo” a fallo y de fallo a “no fallo”. Además de permitir la alternancia entre fallos se puede seleccionar el instante de tiempo en el que ocurre la alternancia. Mediante esta actualización, la simulación de los ciberataques se desarrolla de manera óptima, emulando la secuencia de ciberataque propuesta en el artículo de la Universidad de Burdeos que se ha citado anteriormente.

El bloque tiene dos entradas, la primera es la potencia y la segunda es el tiempo de simulación, de modo que, si en un instante de tiempo concreto, la simulación está en un intervalo de fallo, el bloque devuelve "0" y, si está en un intervalo de no fallo, entonces devuelve el valor de la primera entrada (la potencia).

A continuación, en la Figura 8.4 se muestra el bloque mencionado anteriormente:

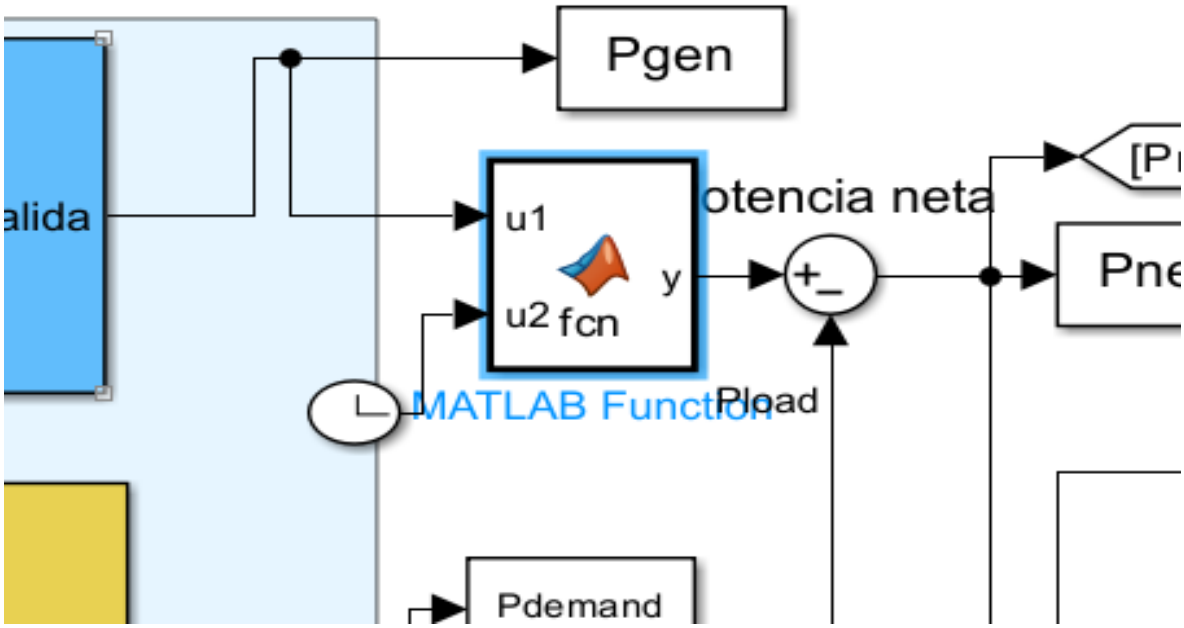

Figura 8.4: *Bloque función de Matlab® programada para crear la secuencia de fallos que simularían un ciberataque.*

```
1   function y = fcn(u1,u2)
2      % Vector de ejemplo con la secuencia de "fallos"
3      % un "1" significa presencia de fallo y un "0" que no lo hay
4      % La primera columna de la variable representa los instantes de tiempo
5      % en los que se producen los cambios fallo/no fallo y viceversa
6      fallos = [0 0; 1000 1; 1010 0; 1020 1; 1030 0; 1050 1; 1100 0];
7
8      z = fallos(find(u2>=fallos(:,1),1,'last'),2);
9      if z == 0
10         y = u1;
11     else
12         y = 0;
13     end
14
```

Figura 8.5: *Líneas de código que componen el bloque función de Matlab®.*

En la Figura 8.4 se puede ver el bloque función de Matlab® el cual permite programar la secuencia de fallos tal y como se desee y en la Figura 8.5 se tiene el desarrollo del código creado para configurar el supuesto ciberataque.

El código mostrado en la Figura 8.5 basa su funcionamiento en la variable *fallos*, la cual permite seleccionar el instante de tiempo en el que se quiere provocar el fallo, además de indicar si se está en presencia de fallo o no. Añadir un *1* a la derecha del tiempo, indicaría que a partir de ese instante de tiempo se encontraría fallando el sistema. De la misma manera, añadiendo un *0* se tendría ausencia de fallo de modo que el funcionamiento del sistema sería el correcto.

Asociado a la existencia o no de fallo, se programa el bloque de modo que si se está en presencia de fallo la potencia de salida del bloque sería nula (tal y como debe de ocurrir en presencia de fallo) y en caso de no existencia de fallo, la potencia de salida del bloque sería la potencia de entrada al bloque.

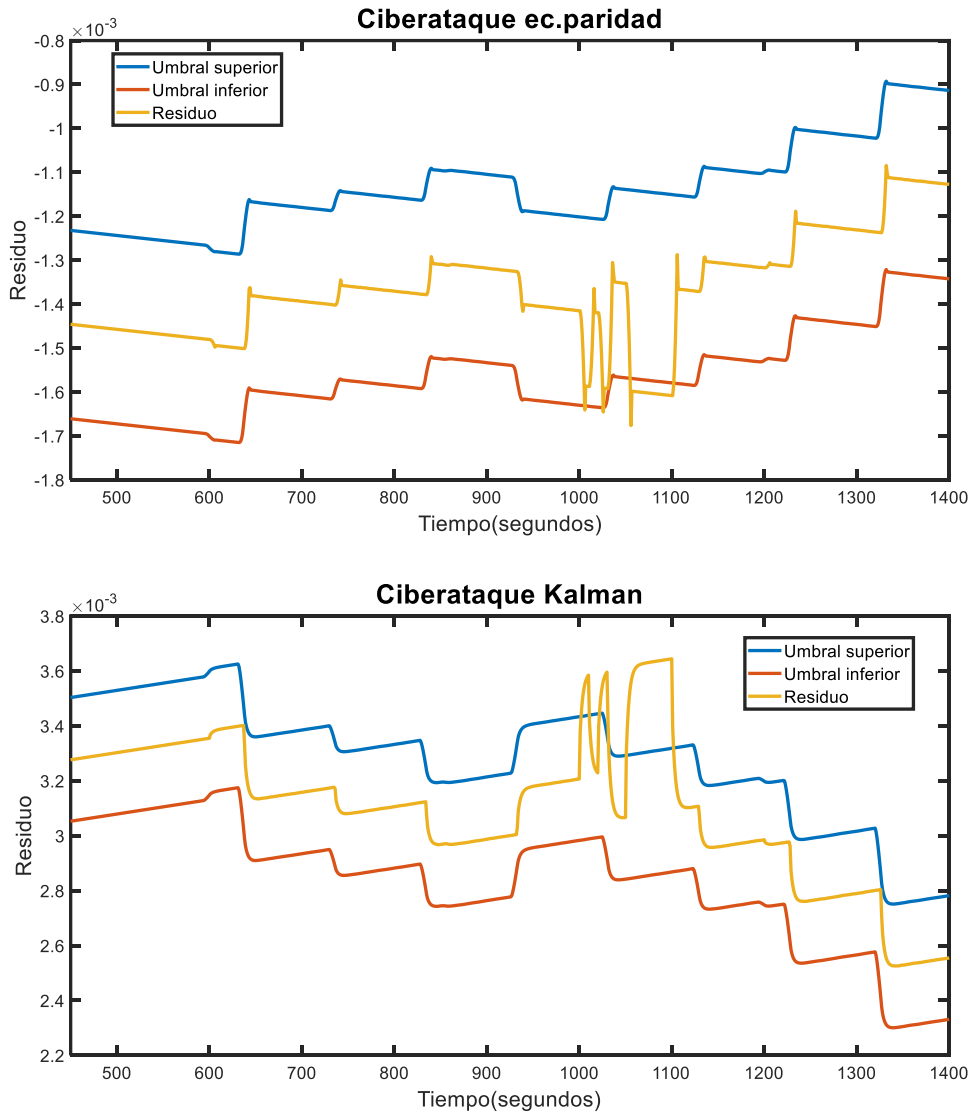

Figura 8.6: *Valores de residuos calculados para un perfil de ciberataque mediante filtro de Kalman y ecuaciones de paridad.*

En la Figura 8.6 se representa el cálculo de residuos ante un perfil de ciberataque diseñado en la Figura 8.5 tanto para el método de ecuaciones de paridad como para filtro de Kalman, así como los umbrales superiores e inferiores.

Para la detección de ciberataques, se ha creado una configuración de modo que, a partir de un número determinado de fallos ocurridos desde el inicio de la simulación, se consideraría que el elemento en cuestión está siendo objetivo de un ciberataque. Como puede observarse, en la gráfica inferior (Figura 8.7), en el instante t = 1050 s se detecta el ciberataque y por tanto la variable binaria "Existencia de ciberataque" toma el valor 1. Se muestra en este caso la variable binaria asociada al perfil de ciberataque mediante filtro de Kalman.

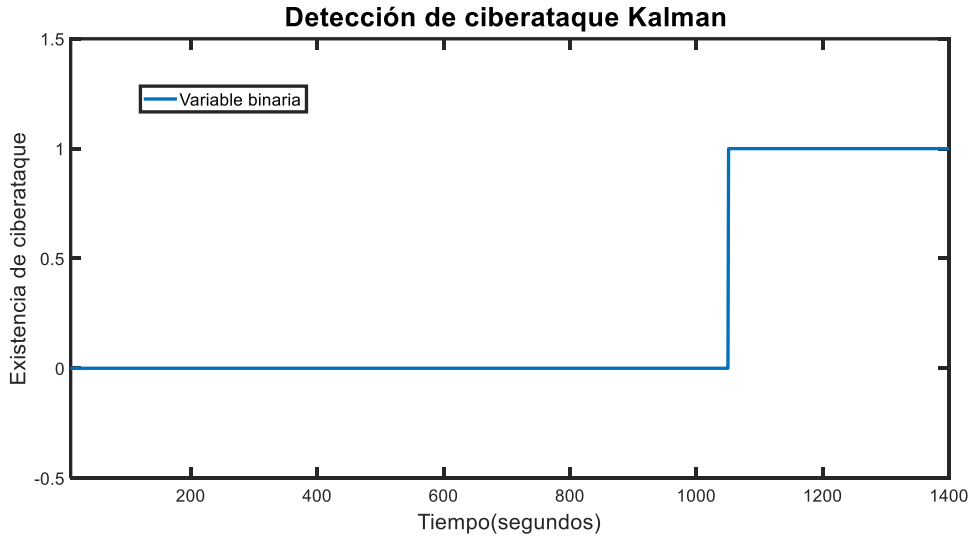

Figura 8.7: *Representación de la variable binaria asociada al perfil de ciberataque mediante filtro de Kalman.*

8.1 Ciberataque a la batería de plomo-ácido

El ciberataque a la batería de plomo-ácido se ha programado para que ocurra en el periodo de tiempo entre $t = 1000\ s$ y $t = 1300\ s$. La frecuencia del ciberataque o el número de fallos ocurridos en ese periodo de tiempo que se asocian a la existencia de ciberataque ha sido seleccionado de manera aleatoria de modo que el "ciberatacante" no ha seguido un patrón definido en el tiempo. La manera de provocar el ciberataque es mediante el bloque función de MatLab® anteriormente explicado, en el intervalo de tiempo definido, se programan al azar una secuencia de fallos y no fallos que alteran la entrada de potencia al componente en cuestión, provocando en consecuencia que el residuo generado sufra variaciones las cuales es necesario detectar para identificar el ciberataque. Se considera que hay existencia de ciberataque a partir de un número de **n=3** fallos detectados.

Las desviaciones estándar, obtenidas a partir de datos históricos, para el cálculo de los umbrales son:

$$\sigma_{paridad}^{pb} = 9.6394 \cdot 10^{-4}\ ;\ \sigma_{Kalman}^{pb} = 0.0010 \qquad (8.1)$$

La media se obtiene en cada tiempo de muestreo, de igual manera para cada elemento, a partir de los datos históricos para una ventana de tiempo $(t + p_1)$ tratándose por tanto de una media móvil siendo $p_1 = 6$.

Los umbrales estocásticos se calculan, tal y como se detalla en el epígrafe 6.4, haciendo uso de la ICDF (*Inverse Cumulative Distribution Function*) de una distribución normal para cada instante de tiempo. El código utilizado en MatLab® (aplicable para el resto de los componentes de la microrred) es (Figura 8.8):

```
for t=1:length(media_pb)

beta_pb(t)=norminv(p,media_pb(t), desviacion_pb);
gamma_pb(t)=norminv(1-p,media_pb(t), desviacion_pb);

beta_pbk(t)=norminv(p,media_pbk(t), desviacion_pbk);
gamma_pbk(t)=norminv(1-p,media_pbk(t), desviacion_pbk);


end
```

Figura 8.8: *Código de MatLab® para el cálculo de los umbrales estocásticos.*

A continuación, se muestran las figuras correspondientes a los ciberataques provocados sobre la batería de plomo-ácido, utilizando los métodos de detección de fallos de ecuaciones de paridad y filtro de Kalman. Los valores de p seleccionados son 0.85, 0.75 y 0.65 respectivamente (Figuras 8.9 a 8.14). Los umbrales superior e inferior se describen mediante las líneas azules y rojas, respectivamente.

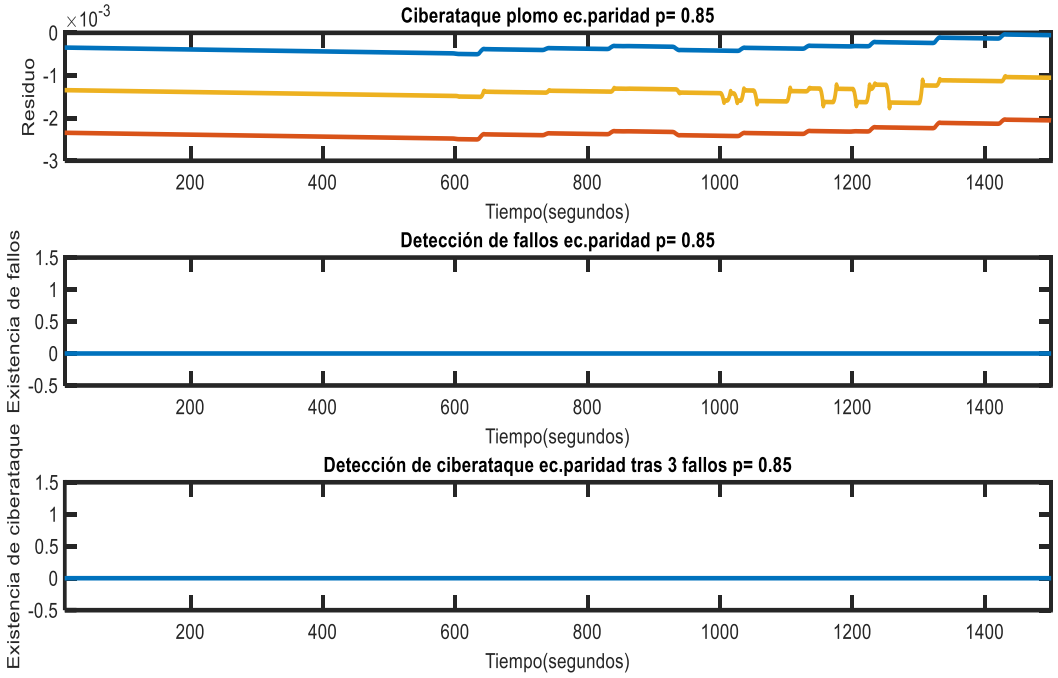

Figura 8.9: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante ec. de paridad con p= 0.85.*

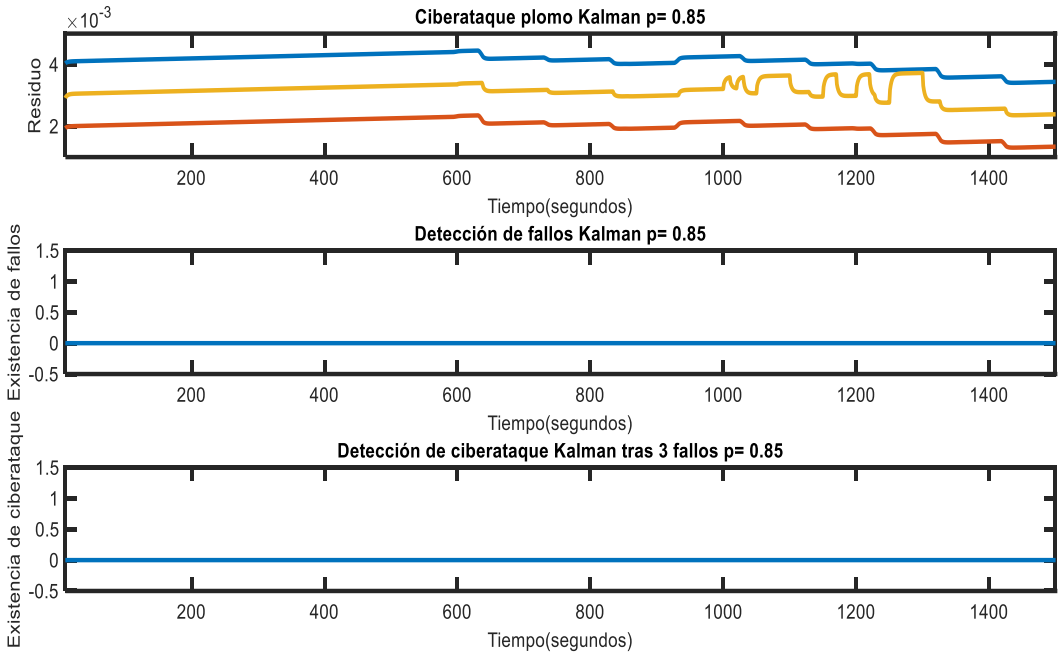

Figura 8.10: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante filtro de Kalman con p= 0.85.*

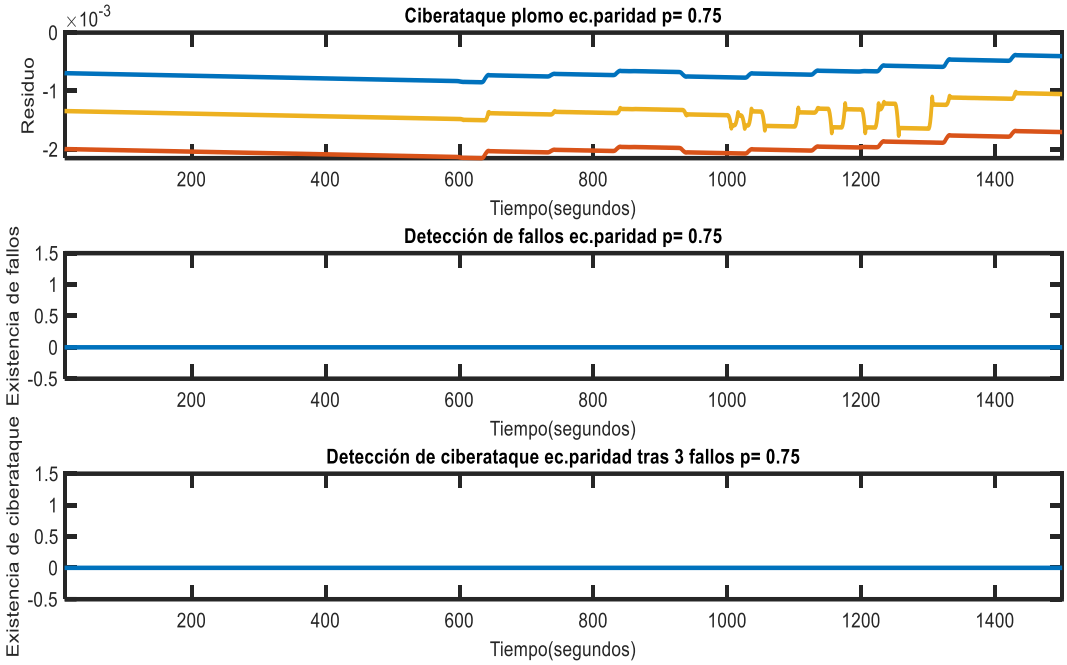

Figura 8.11: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante ec. de paridad con p= 0.75.*

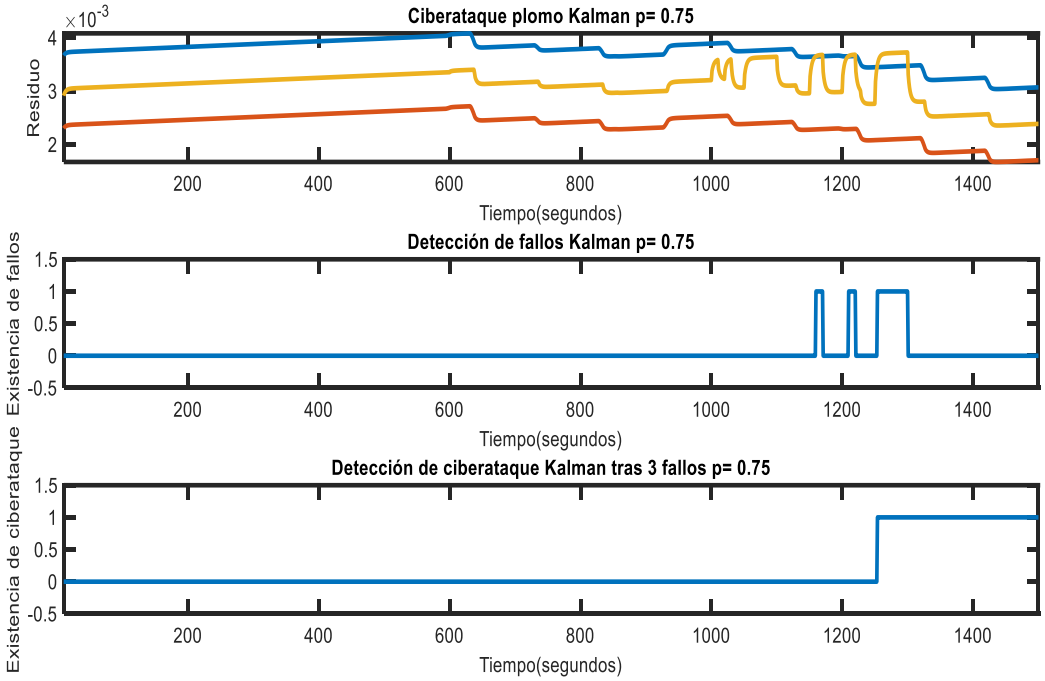

Figura 8.12: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante filtro de Kalman con p= 0.75.*

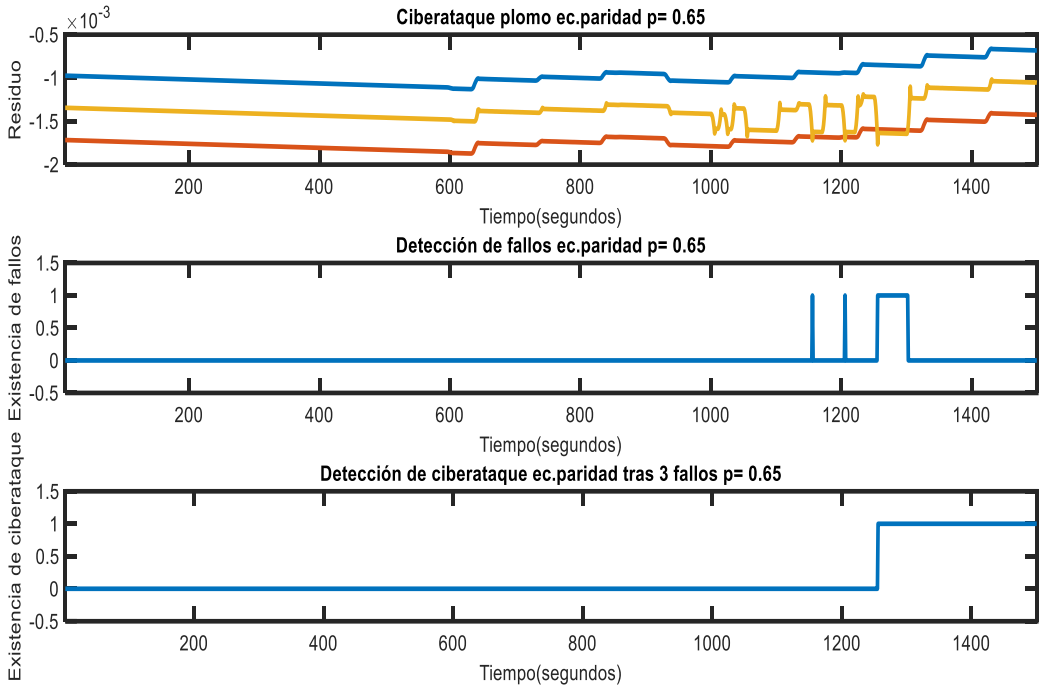

Figura 8.13: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante ec. de paridad con p= 0.75.*

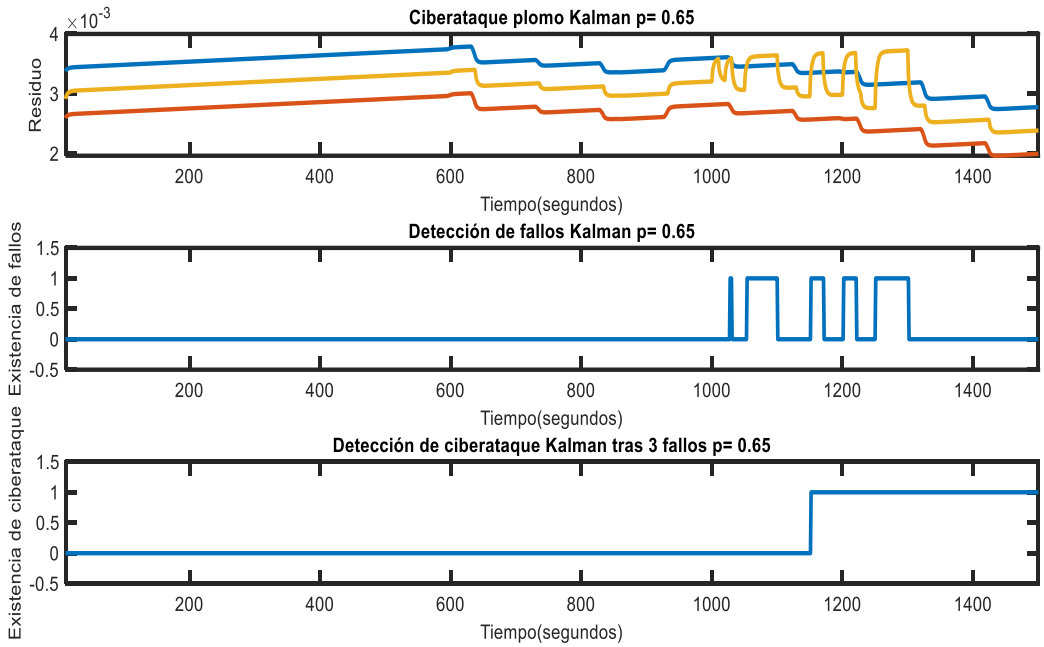

Figura 8.14: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 3 fallos a la batería plomo-ácido mediante filtro de Kalman con p= 0.75.*

En vista a las figuras anteriores, se observa la relevancia de seleccionar un valor de *p* adecuado:

- $\boldsymbol{p = 0.85}$ **:** Se tiene que el algoritmo no es capaz de detectar el ciberataque mediante ninguno de los dos métodos utilizados, de modo que esta situación sería la de falso negativo, dando lugar a que el ciberataque pase desapercibido.
- $\boldsymbol{p = 0.75}$ **:** En este caso, sigue siendo indetectable el ciberataque para el método de ecuaciones de paridad, sin embargo, para el observador basado en el filtro de Kalman se ven reflejados una serie de fallos que son detectados y activan la alarma de detección, la cual ha sido fijada en la detección de tres fallos.
- $\boldsymbol{p = 0.65}$ **:** Finalmente, seleccionando este valor de *p*, se consigue detectar con efectividad la presencia de ciberataque en ambos métodos de detección. A pesar de que el método de ecuaciones de paridad comienza a ser efectivo, es en el método de filtro de Kalman en donde se puede detectar el ciberataque de manera más temprana dando lugar a una capacidad de reacción más temprana.

El hecho de que el filtro de Kalman sea más efectivo que el método basado en las ecuaciones de paridad viene dado porque la magnitud de variación en el caso de filtro de Kalman es mayor, es decir, es más sensible en este caso a los fallos provocados que el método de las ecuaciones de paridad.

8.2 Ciberataque a la batería de litio

El ciberataque a la batería de litio se ha programado para que ocurra en el periodo de tiempo entre $t = 4500\ s$ y $t = 5000\ s$. La frecuencia del ciberataque o el número de fallos ocurridos en ese periodo de tiempo que se asocian a la existencia de ciberataque han sido seleccionados de manera aleatoria de modo que el "ciberatacante" no ha seguido un patrón definido en el tiempo. La manera de provocar el ciberataque es mediante el bloque función de MatLab® anteriormente explicado, en el intervalo de tiempo definido, se programan al azar una secuencia de fallos y no fallos que alteran la entrada de potencia al componente en cuestión, provocando en consecuencia que el residuo generado sufra variaciones las cuales es necesario detectar para identificar el ciberataque. Se considera que hay existencia de ciberataque a partir de un número de **n=4** fallos detectados.

Las desviaciones estándar, obtenidas a partir de datos históricos sin ocurrencia de fallos, para el cálculo de los umbrales son:

$$\sigma_{paridad}^{pb} = 8.9335 \cdot 10^{-4}\ ;\ \sigma_{Kalman}^{pb} = 0.0036 \tag{8.2}$$

A continuación, se muestran las figuras correspondientes a los ciberataques provocados sobre la batería de litio, utilizando los métodos de detección de fallos de ecuaciones de paridad y filtro de Kalman. Los valores de p seleccionados son 0.85, 0.75 y 0.65 respectivamente (Figuras 8.15 a 8.20). Los umbrales superior e inferior se describen mediante las líneas azules y rojas, respectivamente.

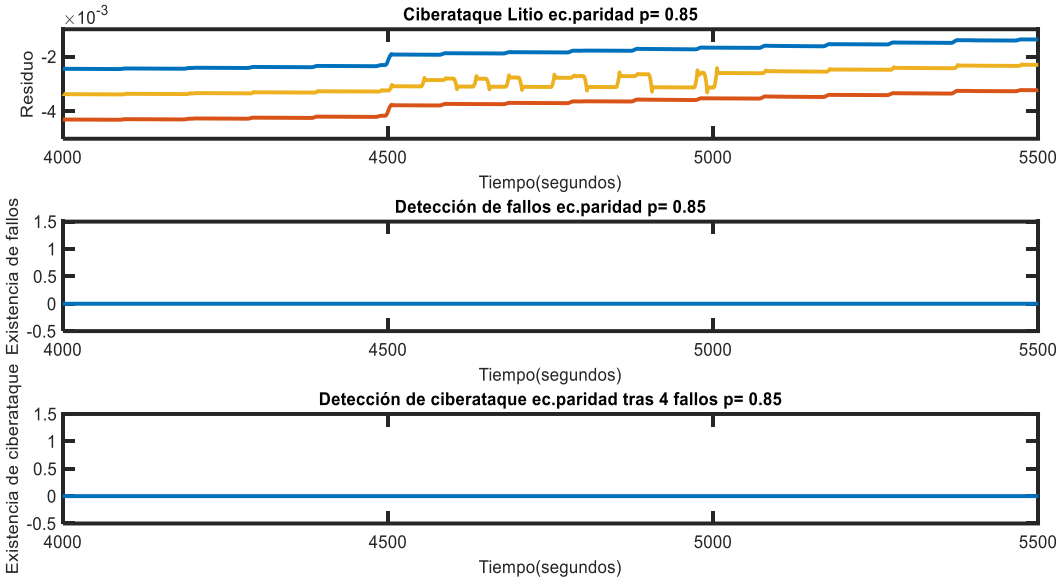

Figura 8.15: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante ec. de paridad con p= 0.85.*

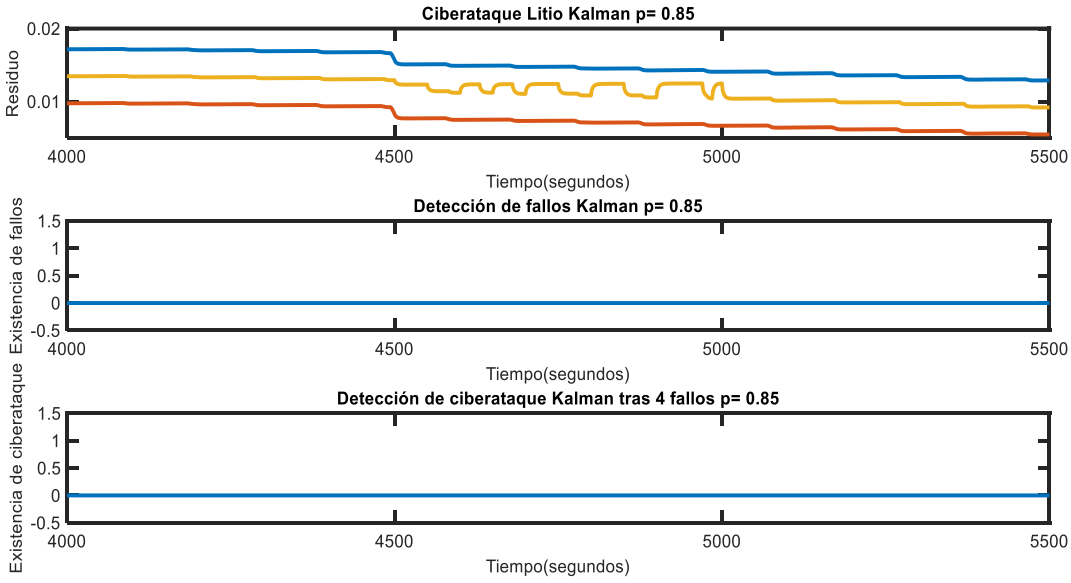

Figura 8.16: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante filtro de Kalman con p= 0.85.*

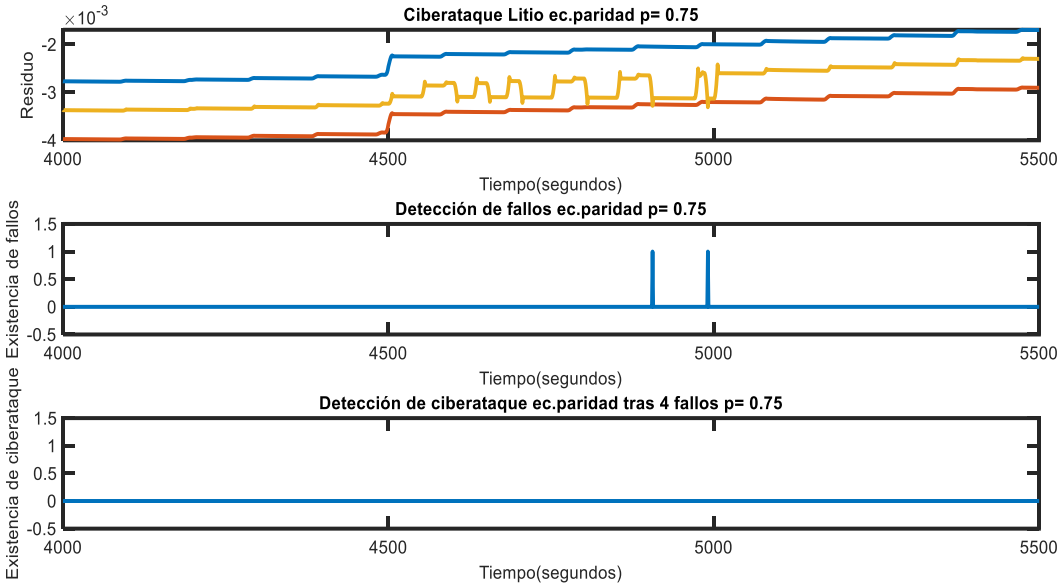

Figura 8.17: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante ec. de paridad con p= 0.75.*

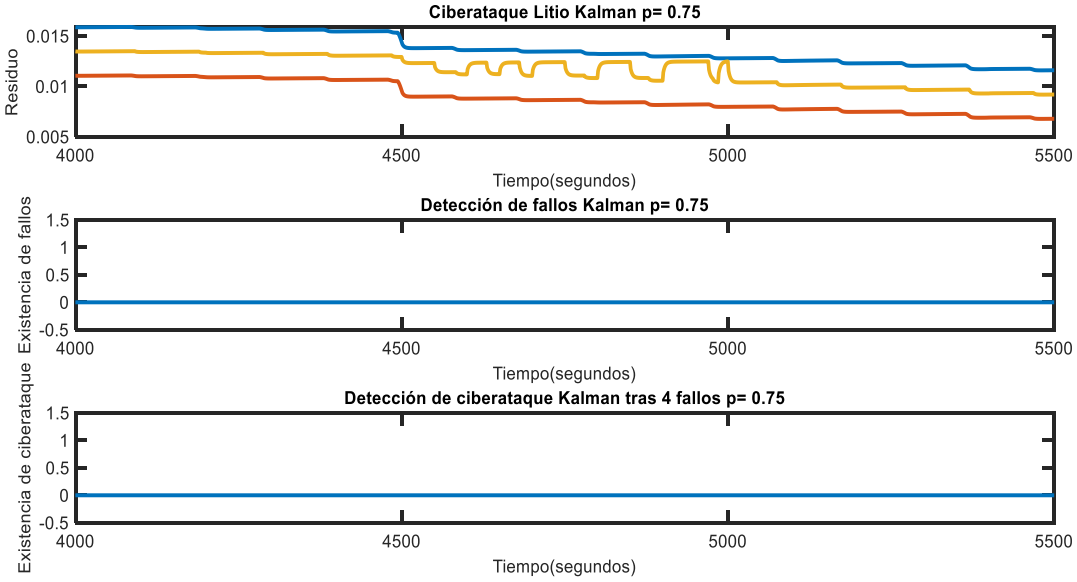

Figura 8.18: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante filtro de Kalman con p= 0.75.*

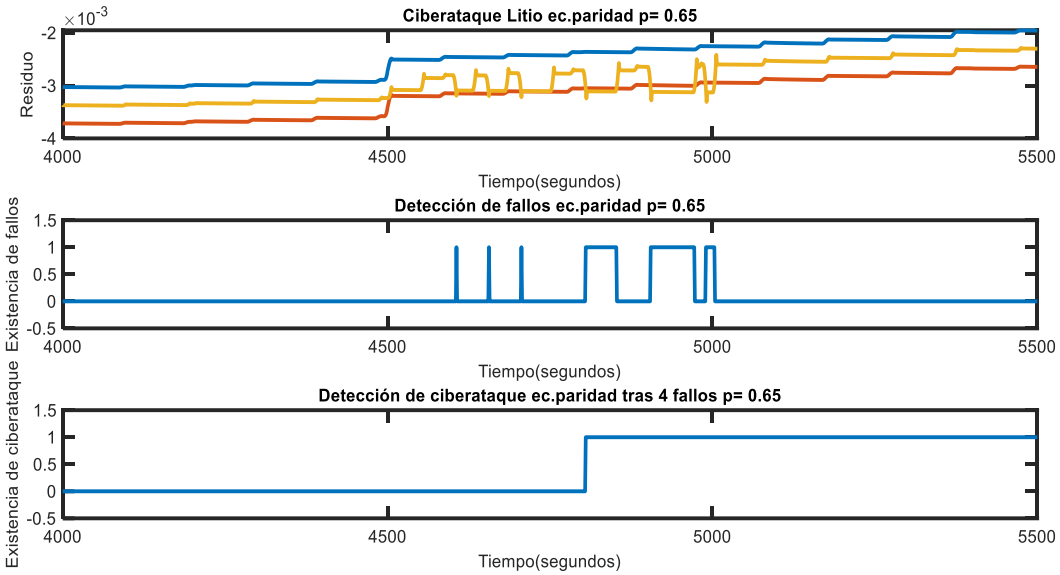

Figura 8.19: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante ec. de paridad con p= 0.65.*

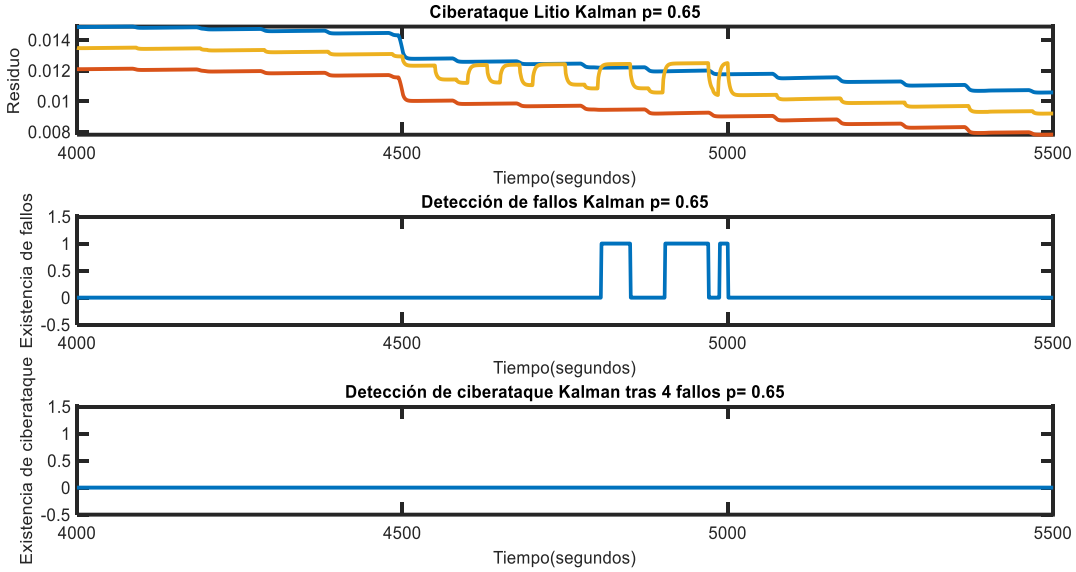

Figura 8.20: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 4 fallos a la batería de litio mediante filtro de Kalman con p= 0.65.*

En vista a las figuras anteriores, se observa la relevancia de seleccionar un valor de *p* adecuado:

- $\boldsymbol{p = 0.85}$ **:** Se tiene que el algoritmo no es capaz de detectar el ciberataque mediante ninguno de los dos métodos utilizados, de modo que esta situación sería la de falso negativo, dando lugar a que el ciberataque pase desapercibido.
- $\boldsymbol{p = 0.75}$ **:** En este caso, sigue siendo indetectable el ciberataque para el método de filtro de Kalman, sin embargo, para el método basado en las ecuaciones de paridad se pueden apreciar dos fallos, los cuales, no serían suficientes para sospechar de un posible ciberataque según el criterio establecido en este caso de detección, el cual considera ciberataque a partir del cuarto fallo. Esto se realiza de esta manera ya que podría tratarse de fallos ocurridos en el sistema por cualquier otro motivo. El establecimiento de un número determinado de fallos a partir de los cuales se podría considerar ciberataque se obtendría de datos históricos y de todos aquellos ciberataques que pudieran ir ocurriendo.
- $\boldsymbol{p = 0.65}$ **:** Finalmente, seleccionando este valor de *p*, se consigue detectar con efectividad la presencia de ciberataque con el método de ecuaciones de paridad. El filtro de Kalman detecta la presencia de tres fallos, los cuales para el criterio establecido de n=4, no serían suficientes para la detección del ciberataque. Tal y como se ha comentado antes, el establecimiento de un número concreto a partir del cual se consideraría ciberataque es algo que dependería de los históricos de fallos ocurridos. Con un elevado número de fallos detectados, se tiene una probabilidad muy elevada de que lo que está provocando esa serie de fallos sea un ciberataque, sin embargo, también podría producirse con un número menor de fallos, de modo que cabría la posibilidad de falsos negativos a la hora de la detección de ciberataques.

Como se puede observar en este caso, el método de detección mediante las ecuaciones de paridad es más efectivo que el filtro de Kalman, lo que no ocurría en el caso anterior, en dónde el filtro de Kalman era más efectivo que el método de ecuaciones de paridad.

La variación del residuo durante el periodo de tiempo seleccionado no presenta saltos bruscos más allá de los provocados intencionadamente para simular el ciberataque, sin embargo, en $t = 4500\ s$ se puede observar un salto brusco en el valor del residuo lo que podría llevar a un falso positivo en la detección de un fallo si se selecciona un valor de *p* muy ajustado. Esta apreciación es uno de los problemas fundamentales que presentan estos métodos de detección, en donde hay que seleccionar con un acertado criterio el valor *p* asociado a los umbrales estocásticos, de manera que se pueda detectar con acierto el posible ciberataque sin llegar a tener la presencia de falsos positivos.

8.3 Ciberataque al supercondensador

El ciberataque a la batería de litio se ha programado para que ocurra en el mismo periodo de tiempo que se producía el ciberataque a la batería de litio, entre $t = 4500\ s$ y $t = 5000\ s$. La frecuencia del ciberataque o el número de fallos ocurridos en ese periodo de tiempo asociados a la existencia de ciberataque, han sido seleccionados de manera aleatoria de modo que el "ciberatacante" no ha seguido un patrón definido en el tiempo. La manera de provocar el ciberataque es mediante el bloque función de MatLab® anteriormente explicado, en el intervalo de tiempo definido, se programan al azar una secuencia de fallos y no fallos que alteran la entrada de potencia al componente en cuestión, provocando en consecuencia que el residuo generado sufra variaciones las cuales es necesario detectar para identificar el ciberataque. Se considera que hay existencia de ciberataque a partir de un número de **n=5** fallos detectados.

Las desviaciones estándar, obtenidas a partir de datos históricos sin ocurrencia de fallos, para el cálculo de los umbrales son:

$$\sigma_{paridad}^{pb} = 2.8737 \cdot 10^{-4}\ ;\ \sigma_{Kalman}^{pb} = 9.1122 \cdot 10^{-4} \tag{8.2}$$

A continuación, se muestran las figuras correspondientes a los ciberataques provocados sobre la batería de litio, utilizando los métodos de detección de fallos de ecuaciones de paridad y filtro de Kalman. Los valores de p seleccionados son 0.85, 0.75 y 0.65 respectivamente (Figuras 8.21 a 8.26). Los umbrales superior e inferior se describen mediante las líneas azules y rojas, respectivamente.

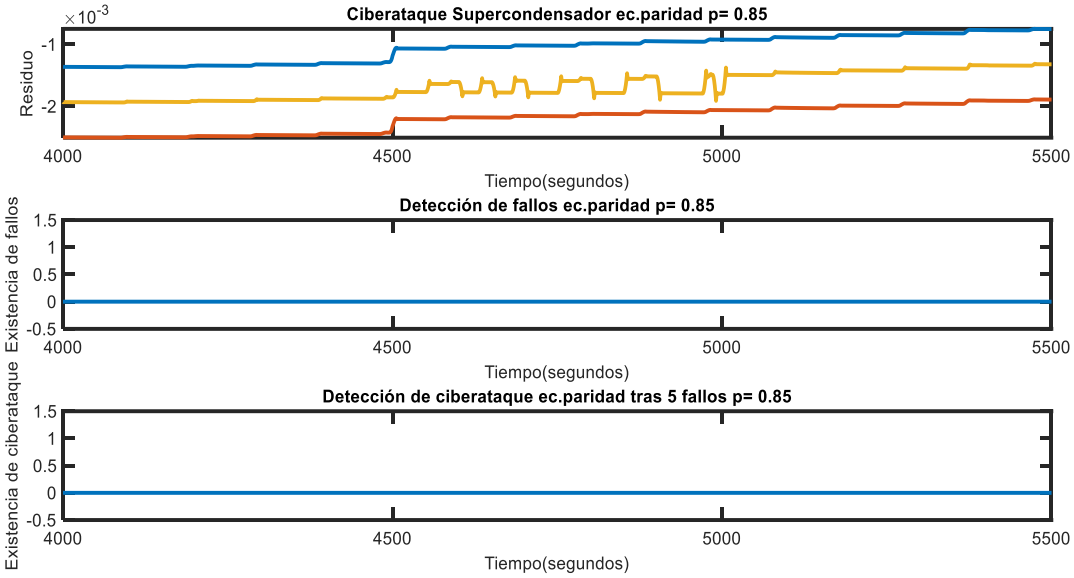

Figura 8.21: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante ec. de paridad con p= 0.85.*

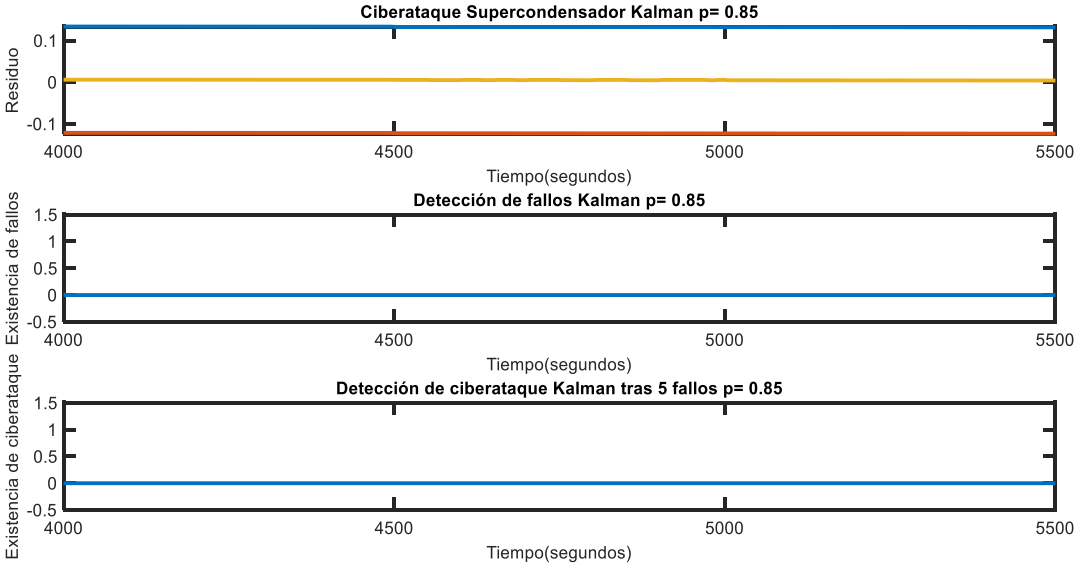

Figura 8.22: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante filtro de Kalman con p= 0.85.*

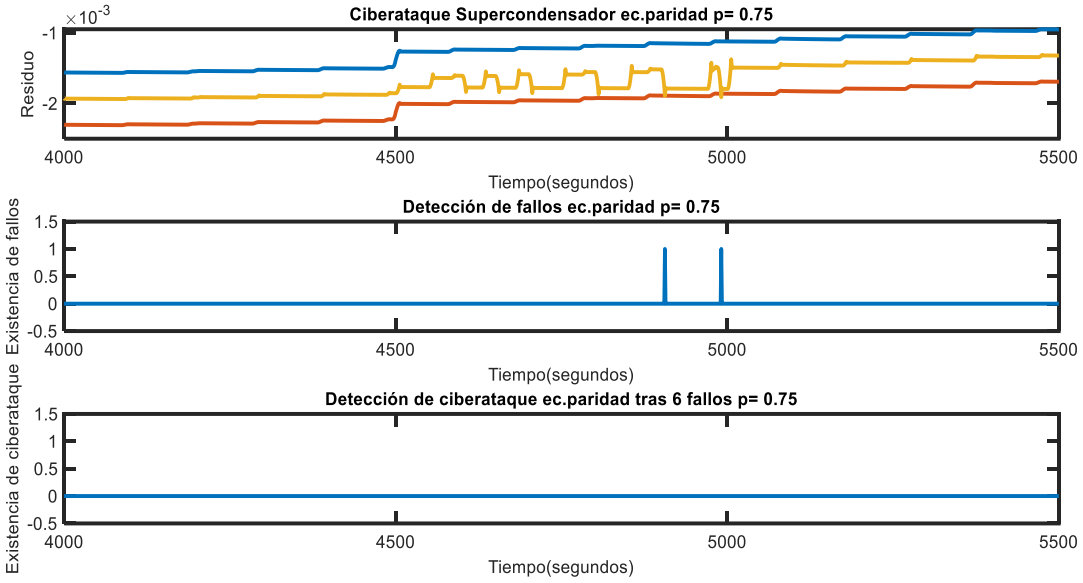

Figura 8.23: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante ec. de paridad con p= 0.75.*

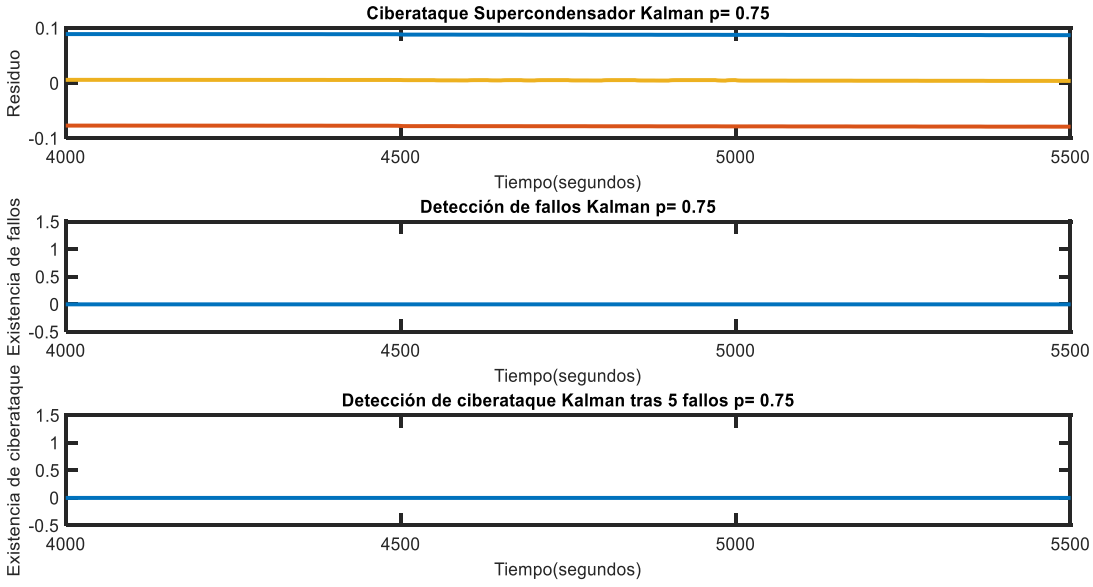

Figura 8.24: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante filtro de Kalman con p= 0.75.*

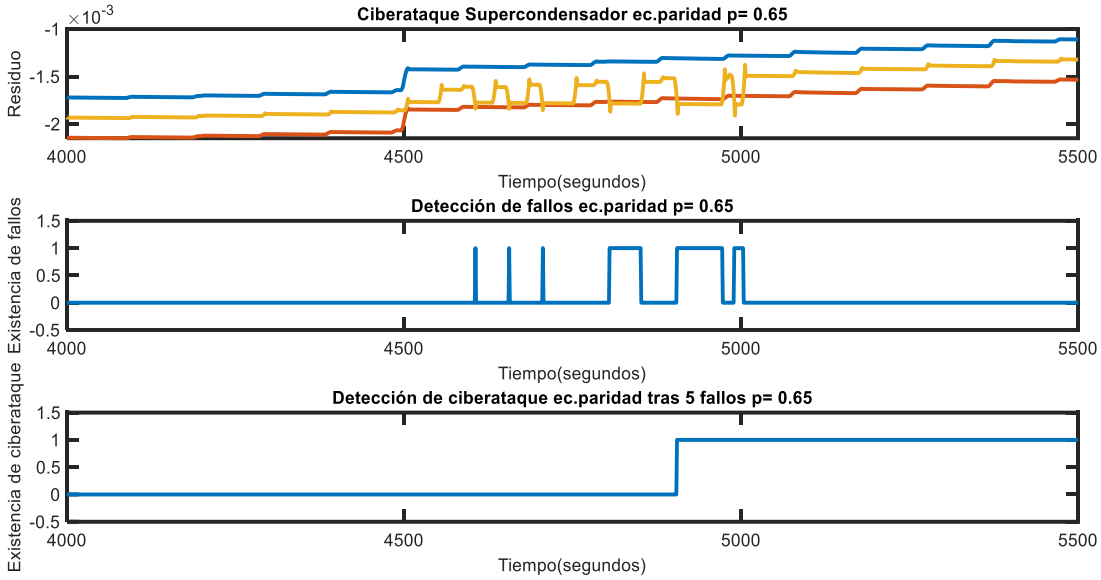

Figura 8.25: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante ec. de paridad con p= 0.65.*

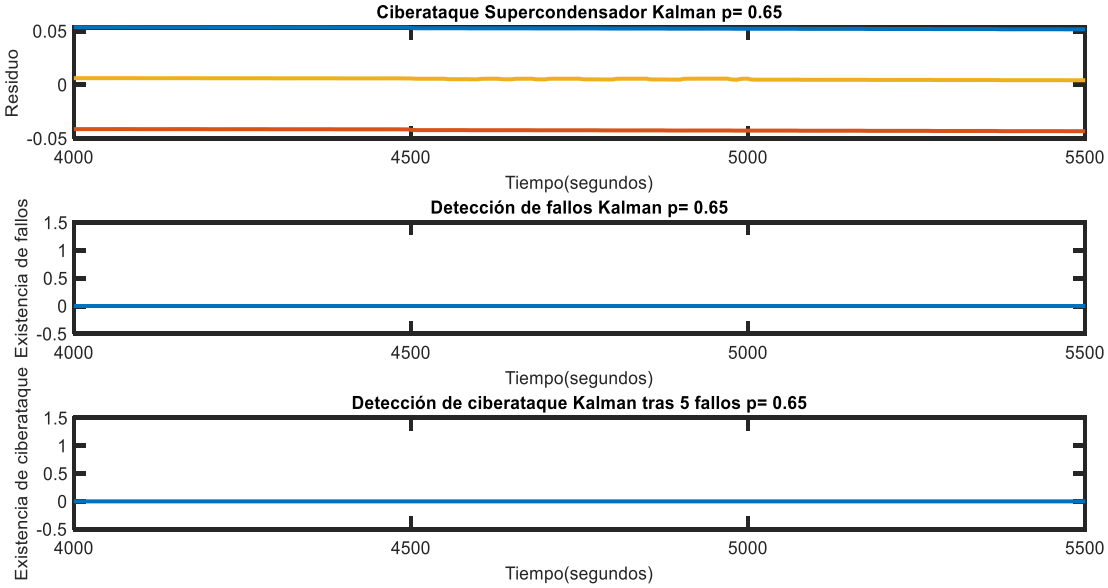

Figura 8.26: *Simulación de ciberataque, detección de fallos y detección de ciberataque tras 5 fallos al supercondensador mediante filtro de Kalman con p= 0.65.*

En vista a las figuras anteriores, se observa la relevancia de seleccionar un valor de *p* adecuado:

- $\boldsymbol{p = 0.85}$**:** Se tiene que el algoritmo no es capaz de detectar el ciberataque mediante ninguno de los dos métodos utilizados, de modo que esta situación sería la de falso negativo, dando lugar a que el ciberataque pase desapercibido.
- $\boldsymbol{p = 0.75}$**:** En este caso, sigue siendo indetectable el ciberataque para el método de filtro de Kalman, sin embargo, para el método basado en las ecuaciones de paridad se pueden apreciar dos fallos, los cuales, no serían suficientes para sospechar de un posible ciberataque habiendo establecido el criterio de detección a partir del quinto fallo.
- $\boldsymbol{p = 0.65}$**:** Finalmente, seleccionando este valor de *p*, se consigue detectar con efectividad la presencia de ciberataque en el método de ecuaciones de paridad.

  En este caso, el método de filtro de Kalman presenta unas variaciones en el residuo muy pequeñas que no llegan a alcanzar los valores de los umbrales seleccionados de modo que para este caso en particular el filtro de Kalman no detecta el ciberataque, la solución básica sería seguir reduciendo aún más el umbral, pero podría conllevar a la aparición de numerosos falsos positivos en instantes de tiempo en los que no estaría ocurriendo el ciberataque. Es por lo ocurrido anteriormente que es de vital importancia disponer de al menos dos métodos de detección de ciberataques de modo que en caso de que uno no detecte se disponga de otro para contrastar el resultado obtenido.

9 Conclusiones

El desarrollo de este trabajo parte de la problemática actual de todo lo relacionado con la seguridad informática, dado que cada día se informatizan y automatizan numerosos procesos los cuales no están exentos de ser víctimas de ataques de interferencia ilícita.

Es de lo anteriormente expuesto de donde nace el concepto de "ciberataque" y posteriormente la respuesta a ello que sería la ciberseguridad. La ciberseguridad es un mundo extremadamente amplio que abarca multitud de facetas diferentes dada la complejidad y la gran variedad de posibilidades mediante las cuales los ciberdelincuentes pueden realizar los actos ilícitos contra los sistemas informáticos.

En este proyecto, se parte de un modelo de microrred proporcionado, al cual, se le han aplicado las modificaciones necesarias para poder realizar simulaciones que emulen el proceso de un ciberataque y las medidas para detectar o al menos tener sospechas de que en efecto se podría estar produciendo.

9.1 Rendimiento del modelo predictivo aplicado a ciberataques

Para la detección de posibles ciberataques, se han utilizado dos métodos distintos, pero con la misma finalidad, la detección del ciberataque mediante un residuo generado. De modo que si se superan unos umbrales establecidos se entiende que el funcionamiento no sería el adecuado y por lo tanto se podría tener sospecha de que un ciberataque pudiera estar ocurriendo.

Lo que diferencia a ambos métodos es la manera de calcular el residuo a partir de las entradas, que son comunes para ambos algoritmos también, a excepción de las variables estadísticas que el observador basado en el filtro de Kalman implementa para una mejor estimación, mientras que las ecuaciones de paridad utilizan únicamente las entradas y salidas del sistema.

En cuanto a los resultados obtenidos, existen factores que determinan la eficiencia de los algoritmos, los cuales son fundamentales para tener en cuenta la correcta detección de los ciberataques, algunos importantes como:

- **Valor de p seleccionado:** Es fundamental seleccionar un valor de p idóneo que permita la correcta detección de los ciberataques sin causar confusiones mediante la detección de falsos positivos ajustando en exceso su valor, o empleando umbrales demasiado grandes de forma que puedan llegar a pasar desapercibidos los posibles ciberataques. La solución correcta vendría determinada por un historial elevado de casos, que permita ir ajustando poco a poco y de manera precisa el rango de valores de p adecuados.

- **Correcto funcionamiento del modelo:** Dado que el funcionamiento de las simulaciones está fundamentado en la utilización de algoritmos basado en modelo, el correcto ajuste de estos es fundamental para que proporcionen unas salidas con valores que se acerquen a la realidad del proceso. No tendría ningún sentido este método de detección de ciberataques si los resultados que están proporcionando los modelos tienen desajustes de tal magnitud que no se correspondan con la realidad.
- **Momento del ciberataque:** Durante la realización de las simulaciones se pudo comprobar que la correcta detección de las cadenas de fallos en serie (posibles ciberataques) tenían una probabilidad mucho más elevada de detectarse si el elemento en cuestión se encontraba en funcionamiento, de modo que si el elemento se encontraba en fase de reposo (provee potencia con menor intensidad) la detección se complicaba.

9.2 Líneas futuras

Las futuras líneas de investigación podrían enfocarse en la mejora de adaptación de los algoritmos a las tácticas cambiantes de los ciberdelincuentes. Esto implica el desarrollo de técnicas de aprendizaje automático que sean capaces de ajustarse dinámicamente a nuevos patrones de ataque y que puedan aprender de manera continua a partir de datos en tiempo real.

En vista a lo expuesto en las líneas anteriores, la inteligencia artificial podría ampliar la identificación de patrones más complejos y la correlación de información a gran escala. Mediante el uso de la inteligencia artificial y una base de datos de ciberataques a lo largo de la historia, se podría conseguir llegar a una precisión adecuada para que se detecten de manera correcta los ciberataques a pesar de las posibles variaciones que los ciberdelincuentes puedan ir añadiendo.

Como ejemplo de mejora al modelo propuesto se tiene la posibilidad de monitorizar la diferencia entre los umbrales y el residuo generado, de modo que cuando experimente caídas por debajo de otro umbral seleccionado se puedan tener sospechas de la existencia de ciberataque, todo esto sin necesidad de que se sobrepasen los umbrales iniciales, aunque sí que deberían de estar muy cerca de ser sobrepasados.

A continuación (Figura 9.1) se muestra un perfil de ciberataque que está muy cerca de sobrepasar los umbrales pero no llega a sobrepasarlos:

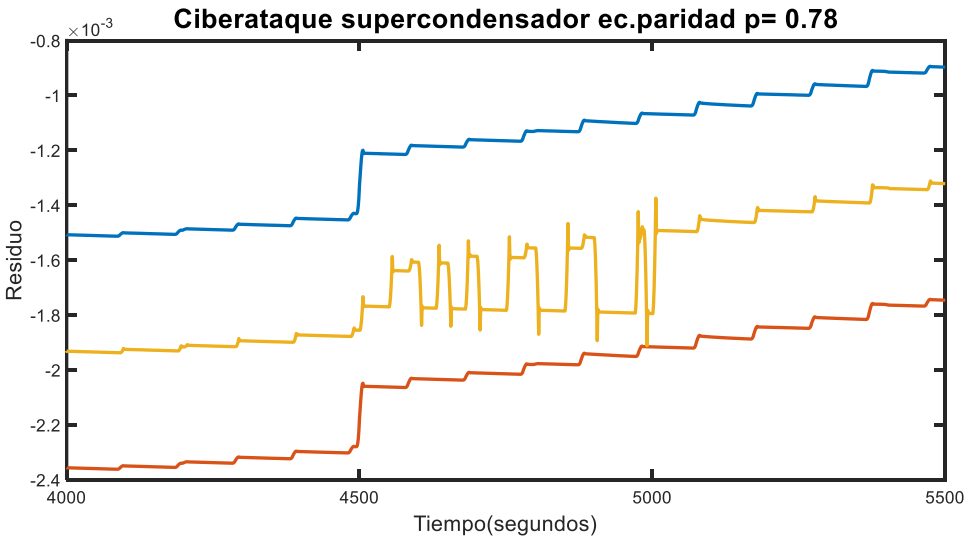

Figura 9.1: *Valores residuales y umbrales estocásticos mediante ec. de paridad representados en el ciberataque al supercondensador con p=0.78.*

Como se puede apreciar, para el valor de *p=0.7*8 se consigue estar muy cerca de los umbrales, pero no se llegan a sobrepasar, para ello se podría diseñar la siguiente solución (Figura 9.2):

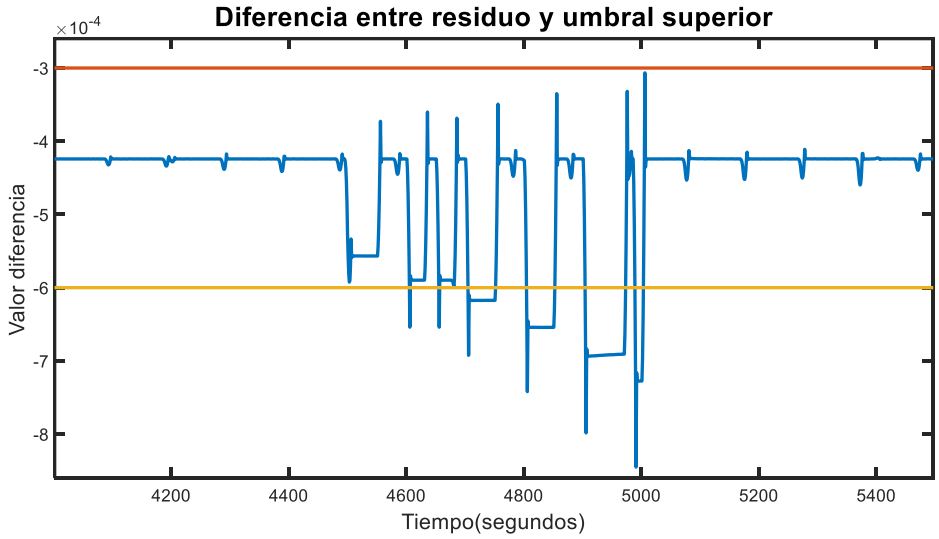

Figura 9.2: *Diferencia entre residuo generado y umbral superior para un ciberataque al supercondensador para un valor de p=0.78 mediante ec. de paridad.*

En la figura mostrada se muestra el valor diferencial entre el residuo generado y el umbral superior (se podría hacer con el umbral inferior también). Se puede observar una clara alteración en el valor mientras se produce el ciberataque en comparación con su funcionamiento correcto. Se han posicionado como ejemplo unos nuevos umbrales superior e inferior que podrían servir para detectar la posibilidad de ciberataque.

Esto serviría como un método auxiliar o adicional de detección monitorizando a su vez los valores diferenciales entre umbrales y residuo, teniendo en cuenta que habría que ajustarlos correctamente para evitar la detección de falsos positivos.

Como método de defensa ante la repentina detección de un posible ciberataque, se tendría la posibilidad de desactivar el componente atacado en cuestión. Para ello se procede a realizar una reconfiguración de modo que una vez detectado el ciberataque se anule la utilización del componente. Para ejemplificarlo se va a suponer que en $t = 1300\ s$ se ha detectado la posible presencia de un ciberataque a la batería de plomo-ácido mediante alguno de los métodos citados anteriormente y de manera automática se inhabilita el funcionamiento de la batería para proteger el sistema.

Para mostrar el efecto de la reconfiguración se tienen las siguientes imágenes:

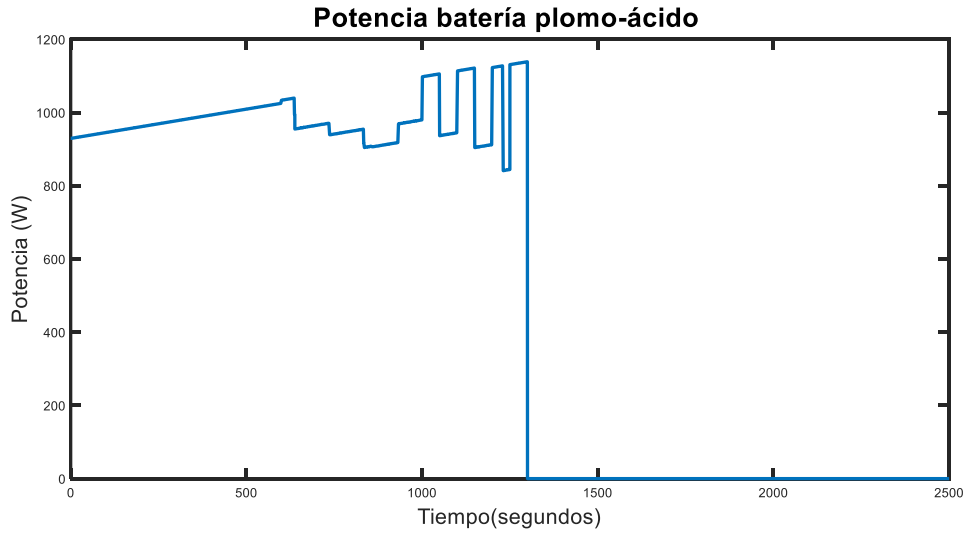

Figura 9.3: *Evolución de la potencia de la batería plomo-ácido mientras se produce un ciberataque detectado en t = 1300 s.*

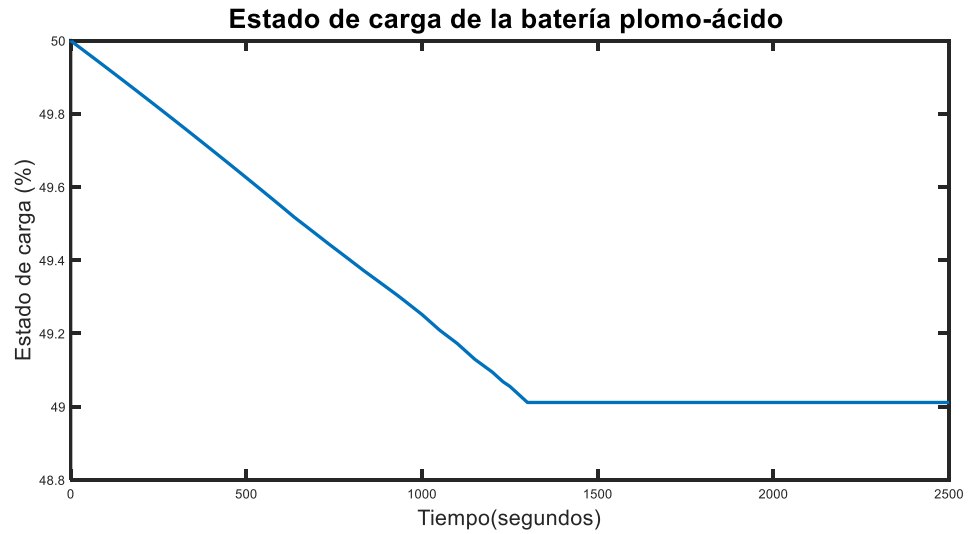

Figura 9.4: *Evolución del estado de carga de la batería plomo-ácido mientras se produce un ciberataque detectado en t = 1300 s.*

En la Figura 9.3 se muestra la evolución de la potencia de la batería plomo-ácido ante la presencia de un ciberataque detectado en $t = 1300\ s$. Como se puede observar, a partir del segundo 1300, el valor de potencia de la batería se sitúa en 0, indicando que se ha desactivado el elemento en cuestión.

En la Figura 9.4 se muestra la evolución del estado de carga de la batería plomo-ácido ante la presencia de un ciberataque detectado en $t = 1300\ s$. En este caso se aprecia que a partir del segundo 1300, el estado de carga de la batería se mantiene constante con el valor que tenía en el mismo instante que es detectado el ciberataque. Esto indica que la batería no está siendo utilizada como efecto de la desactivación para proteger el elemento del ciberataque.

Las mejoras explicadas anteriormente, junto con la utilización de la inteligencia artificial, permitirán mejorar la protección ante interferencias ilícitas de carácter informático.

Índice de Figuras

Bibliografía

[1] CEDEA. https://www.inta.es/INTA/es/quienes-somos/historia/el-arenosillo/.

[2] «CENER,» https://www.cener.com/introduccion-a-las-microrredes/.

[3] C. Bordons, F. García-Torres y M. A.Ridao, *Model predictive control of microgrids*, 2019.

[4] E. F. Camacho y C. Bordons, *Model predictive control*, 2007.

[5] S. R. Moreno, *Gestión de la energía de una microrred mediante control predictivo basado en modelo*, 2019.

[6] MathWorks. https://es.mathworks.com/.

[7] ResearchGate. https://www.researchgate.net/figure/Figura-2-Modelo-general-de-una-microrred-Fuente-Elaboracion-propia-Los-esquemas-de_fig2_313024019.

[8] «GlobalElectricity,» https://globalelectricity.wordpress.com/2014/01/30/componentes-de-las-microrredes-electricas-o-microgrids/.

[9] «Albarenova,» https://albarenova.com/microrredes/.

[10] Zhang, H., Li, P., Zhang, W., & Zhang, J. (2021). *A Review of Cybersecurity Issues and Solutions in Smart Grids and Microgrids*. Energies.

[11] «Energystorage,» https://www.energy-storage.news/battery-systems-for-microgrids-on-three-continents-on-the-way-from-rolls-royce-fluence-duke-energy/.

[12] «ScienceDirect,» https://www.sciencedirect.com/science/article/abs/pii/S0045790622007716#:~:text=Microgrid%20with%20IoT%20for%20energy%20management%20and%20control,-

The%20proposed%20microgrid&text=The%20IoT%2Dbased%20module%20is,low%20demand%20and%20high%20demand..

[13] P. A. F. Gómez, *Automatización de una microrred con doble banco de baterías*, Universidad de Sevilla, 2018.

[14] ThinkMicrogrid. https://www.thinkmicrogrid.org/.

[15] «ScienceDirect,» https://www.sciencedirect.com/science/article/pii/S266654682200009X.

[16] «EvolK,» https://evolk.es/que-es-la-ciberseguridad-perimetral/.

[17] «Xataka,» https://www.xataka.com/historia-tecnologica/no-esto-no-es-un-molino-es-el-primer-sistema-de-telecomunicaciones-un-telegrafo-optico.

[18] Telefónica, «telefonica,» https://www.telefonica.com/es/sala-comunicacion/blog/que-es-un-ciberataque-que-tipos-existen-y-para-que-sirve/.

[19] GrupoIca. https://www.grupoica.com/blog/-/blogs/9-tipos-ciberataque-debes-conocer.

[20] Zhang, H., Li, P., Zhang, W., & Zhang, J. (2021). *A Review of Cybersecurity Issues and Solutions in Smart Grids and Microgrids*. Energies.

[21] Pournaras, E., Xydis, G., Kourtessis, P., & Alcaraz-Calero, J. M. (2020). *Securing microgrids: challenges and opportunities*. IEEE Power and Energy Magazine

[22] «IDDigitalSchool,» Available: https://iddigitalschool.com/bootcamps/que-es-la-criptografia/#:~:text=La%20criptograf%C3%ADa%20en%20ciberseguridad%20es,sistemas%20de%20computadoras%20y%20redes..

[23] «CheckPoint» https://www.checkpoint.com/es/cyber-hub/network-security/what-is-an-intrusion-detection-system-ids/.

[24] «Mobbeel,» https://www.mobbeel.com/blog/que-es-la-autenticacion-multifactor-o-de-2-factores/#:~:text=La%20Autenticaci%C3%B3n%20Multifactor%20(MFA)%20es,a%20una%20cuenta%20o%20plataforma..

[25] «Elperiodicodelaenergia,» https://elperiodicodelaenergia.com/la-ciberseguridad-se-convierte-en-una-prioridad-de-las-electricas-por-el-aumento-de-las-amenazas-durante-la-covid-19/.

[26] «Antiun,» https://www.antiun.com/firewall/.

[27] I. N. d. Ciberseguridad, «incibe,» https://www.incibe.es/incibe/formacion.

[28] «ciberseguridad,» https://ciberseguridad.com/guias/nuevas-tecnologias/computacion-cuantica/.

[29] I. N. d. Ciberseguridad, «incibe,» https://www.incibe.es/incibe/sala-de-prensa/incibe-firma-cinco-convenios-para-formar-en-ciberseguridad-mas-de-1700.

[30] INTA. https://www.inta.es/INTA/es/index.html.

[31] R. Isserman, *Fault-diagnosis system: An introduction from fault detection to fault tolerance*, 2005.

[32] T. V. Berbesi, *Aplicación de técnicas robustas para detección y diagnóstico de fallos*, Universidad de Valladolid: Ph.D. thesis, 2012.

[33] V. Puig, J. Quevedo, T. Escobet, B. Morcego y C. Ocampo, *Control tolerante a fallos (parte I): Fundamentos y diagnóstico de fallos*, 2004.

[34] J. Chen, *Model-based methods for fault diagnosis*, 1995.

[35] J. J. M. Quintero, *Control predictivo tolerante a fallos aplicado a sistemas de energía*, Universidad de Sevilla, 2021.

[36] L. F. B. Quintana y L. J. d. M. González, *Diagnóstico de fallos basado en el modelo de la planta*, Universidad de León, 2003.

[37] U. Wang y J. Sarangapani, «*Cyber-attack detection and mitigation in control systems: A model predictive control approach*» *Journal of Process Control,* 2016.

[38] J. d. D. P. Guevara, *Detección de fallos de una microrred mediante algoritmos basados en modelo,* Universidad de Sevilla, 2022.

[39] B.Auboin, A.Perodou, C.Combastel, A.Zolghadri, *Resilient tube-based MPC for Cyber-Physical Systems Under DoS Attacks*, Universidad de Burdeos 2022.

Anexo: Códigos de MatLab®

Se adjunta a continuación los códigos de MatLab® utilizados para la elaboración de las simulaciones de los ciberataques, detección de fallos y detección de ciberataques.

```
clear;
close all;

load("residuos24h.mat"); load("Res_ec_paridad_ciber2.mat");
load("residuos24h_kalman.mat"); load("Res_ec_kalman_ciber2.mat");
r_Ppb=r_nofault.data(:,1); %fallo en t=8000s
r_Ppbf=r.data(:,1);
r_Ppbk=r_kalmannofault.data(:,1);
r_Ppbkf=r_kalman.data(:,1);
p=0.85;

desviacion_pb=std(r_Ppb);
desviacion_pbk=std(r_Ppbk);


%calculo medias

for i=1:(length(r_Ppb)-6)
    media_pb(i)=mean(r_Ppb(i:(6+i)));
    media_pbk(i)=mean(r_Ppbk(i:(6+i)));
end


for t=1:length(media_pb)

beta_pb(t)=norminv(p,media_pb(t), desviacion_pb);
gamma_pb(t)=norminv(1-p,media_pb(t), desviacion_pb);

beta_pbk(t)=norminv(p,media_pbk(t), desviacion_pbk);
gamma_pbk(t)=norminv(1-p,media_pbk(t), desviacion_pbk);


end

% Número de fallos para presencia de ciberataque
ref_num_fallos = 3;

cuenta_pb = 0;
for t=1:length(r_Ppbf)
    if r_Ppbf(t)>beta_pb(t) | r_Ppbf(t)<gamma_pb(t)
        rb_pb(t)=1;
        if t==1
            rb_pb_ant=0;
        else
            rb_pb_ant=rb_pb(t-1);
        end
        if rb_pb(t)==1 && rb_pb_ant==0
            cuenta_pb = cuenta_pb + 1;
        end
    else
        rb_pb(t)=0;
    end
    if cuenta_pb >= ref_num_fallos
        detectado_pb(t) = 1;
    else
```

```
        detectado_pb(t) = 0;
    end
end

cuenta_pbk = 0;
for t=1:length(r_Ppbkf)
    if r_Ppbkf(t)>beta_pbk(t) | r_Ppbkf(t)<gamma_pbk(t)
        rb_pbk(t)=1;
        if t==1
            rb_pbk_ant=0;
        else
            rb_pbk_ant=rb_pbk(t-1);
        end
        if rb_pbk(t)==1 && rb_pbk_ant==0
            cuenta_pbk = cuenta_pbk + 1;
        end
    else
        rb_pbk(t)=0;
    end
    if cuenta_pbk >= ref_num_fallos
        detectado_pbk(t) = 1;
    else
        detectado_pbk(t) = 0;
    end
end

 %
 % figure(1); plot(beta_pb); title(['Ciberataque plomo ec.paridad p= '  num2str(p)]);↙
xlabel('Tiempo(segundos)');ylabel('Residuo'); hold on; plot(gamma_pb); plot(r_Ppbf);↙
xlim([10 1500]);
 % figure(2); plot(beta_pbk); title(['Ciberataque plomo Kalman p= '  num2str(p)]);↙
xlabel('Tiempo(segundos)');ylabel('Residuo'); hold on; plot(gamma_pbk); plot(r_Ppbkf);↙
xlim([10 1500]);
 %
 % figure(3); plot(rb_pb);title(['Detección de fallos ec.paridad p= '  num2str(p)]);↙
xlabel('Tiempo(segundos)');ylabel('Existencia de fallos'); xlim([10 1500]); ylim([-0.5↙
1.5]);
 % figure(4); plot(rb_pbk);title(['Detección de fallos Kalman p= '  num2str(p)]);↙
xlabel('Tiempo(segundos)');ylabel('Existencia de fallos'); xlim([10 1500]); ylim([-0.5↙
1.5]);
 %
 % figure(5); plot(detectado_pb);title(['Detección de ciberataque ec.paridad tras '↙
num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)]); xlabel('Tiempo(segundos)');↙
ylabel('Existencia de ciberataque'); xlim([10 1500]); ylim([-0.5 1.5]);
 % figure(6); plot(detectado_pbk);title(['Detección de ciberataque Kalman tras '↙
num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)]); xlabel('Tiempo(segundos)');↙
ylabel('Existencia de ciberataque'); xlim([10 1500]); ylim([-0.5 1.5]);


% Primera figura con tres subplots
figure(1);
subplot(3, 1, 1);
plot(beta_pb);
hold on;
plot(gamma_pb);
```

```
plot(r_Ppbf);
xlim([10 1500]);
title(['Ciberataque plomo ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
plot(rb_pb);
hold on;
% plot(rb_pbk);
xlim([10 1500]);
ylim([-0.5 1.5]);
title(['Detección de fallos ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
plot(detectado_pb);
hold on;
% plot(detectado_pbk);
xlim([10 1500]);
ylim([-0.5 1.5]);
title(['Detección de ciberataque ec.paridad tras ' num2str(ref_num_fallos) ' fallos '↙
'p= '  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');

% Segunda figura con tres subplots
figure(2);
subplot(3, 1, 1);
plot(beta_pbk);
hold on;
plot(gamma_pbk);
plot(r_Ppbkf);
xlim([10 1500]);
title(['Ciberataque plomo Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
% plot(rb_pb);
% hold on;
plot(rb_pbk);
xlim([10 1500]);
ylim([-0.5 1.5]);
title(['Detección de fallos Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
% plot(detectado_pb);
% hold on;
plot(detectado_pbk);
xlim([10 1500]);
ylim([-0.5 1.5]);
```

```
title(['Detección de ciberataque Kalman tras ' num2str(ref_num_fallos) ' fallos ' 'p=↙
'  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');
```

```
clear;
close all;

load("residuos24h.mat"); load("Res_ec_paridad_ciber_litio.mat");
load("residuos24h_kalman.mat"); load("Res_ec_kalman_ciber_litio.mat");
r_Pli=r_nofault.data(:,2); %Ciberataque entre t=4500 y t=5000
r_Plif=r.data(:,2);
r_Plik=r_kalmannofault.data(:,2);
r_Plikf=r_kalman.data(:,2);
p=0.65;

desviacion_li=std(r_Pli);
desviacion_lik=std(r_Plik);


%calculo medias

for i=1:(length(r_Pli)-6)
    media_li(i)=mean(r_Pli(i:(6+i)));
    media_lik(i)=mean(r_Plik(i:(6+i)));
end


for t=1:length(media_li)

beta_li(t)=norminv(p,media_li(t), desviacion_li);
gamma_li(t)=norminv(1-p,media_li(t), desviacion_li);

beta_lik(t)=norminv(p,media_lik(t), desviacion_lik);
gamma_lik(t)=norminv(1-p,media_lik(t), desviacion_lik);


end

% for t=1:length(r_Plif)
%     if r_Plif(t)>beta_li(t) | r_Plif(t)<gamma_li(t)
%         rb_li(t)=1;
%     else
%         rb_li(t)=0;
%     end
% end
% for t=1:length(r_Plikf)
%     if r_Plikf(t)>beta_lik(t) | r_Plikf(t)<gamma_lik(t)
%         rb_lik(t)=1;
%     else
%         rb_lik(t)=0;
%     end
% end

% Número de fallos para presencia de ciberataque
ref_num_fallos = 4;

cuenta_li = 0;
for t=1:length(r_Plif)
    if r_Plif(t)>beta_li(t) | r_Plif(t)<gamma_li(t)
```

```
        rb_li(t)=1;
        if t==1
            rb_li_ant=0;
        else
            rb_li_ant=rb_li(t-1);
        end
        if rb_li(t)==1 && rb_li_ant==0
            cuenta_li = cuenta_li + 1;
        end
    else
        rb_li(t)=0;
    end
    if cuenta_li >= ref_num_fallos
        detectado_li(t) = 1;
    else
        detectado_li(t) = 0;
    end
end

cuenta_lik = 0;
for t=1:length(r_Plikf)
    if r_Plikf(t)>beta_lik(t) | r_Plikf(t)<gamma_lik(t)
        rb_lik(t)=1;
        if t==1
            rb_lik_ant=0;
        else
            rb_lik_ant=rb_lik(t-1);
        end
        if rb_lik(t)==1 && rb_lik_ant==0
            cuenta_lik = cuenta_lik + 1;
        end
    else
        rb_lik(t)=0;
    end
    if cuenta_lik >= ref_num_fallos
        detectado_lik(t) = 1;
    else
        detectado_lik(t) = 0;
    end
end
```

```
% figure(1); plot(beta_li); title(['Ciberataque litio ec.paridad p= '  num2str↙
(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel↙
('Residuo','FontSize', 20); hold on; plot(gamma_li); plot(r_Plif); xlim([4000 5500]);
% figure(2); plot(beta_lik); title(['Ciberataque litio Kalman p= '  num2str↙
(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel↙
('Residuo','FontSize', 20); hold on; plot(gamma_lik); plot(r_Plikf); xlim([4000↙
5500]);
% figure(3); plot(rb_li); title(['Detección de fallos ec.paridad p= '  num2str↙
(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de↙
```

```
fallos','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);
% figure(4); plot(rb_lik); title(['Detección de fallos Kalman p= '  num2str↙
(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de↙
fallos','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);
% figure(5); plot(detectado_li);title(['Detección de ciberataque ec.paridad tras '↙
num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo↙
(segundos)','FontSize', 20);ylabel('Existencia de ciberataque','FontSize', 20); xlim↙
([4000 5500]); ylim([-0.5 1.5]);
% figure(6); plot(detectado_lik);title(['Detección de ciberataque Kalman tras '↙
num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo↙
(segundos)','FontSize', 20);ylabel('Existencia de ciberataque','FontSize', 20); xlim↙
([4000 5500]); ylim([-0.5 1.5]);


% Primera figura con tres subplots
figure(1);
subplot(3, 1, 1);
plot(beta_li);
hold on;
plot(gamma_li);
plot(r_Plif);
xlim([4000 5500]);
title(['Ciberataque Litio ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
plot(rb_li);
hold on;
% plot(rb_pbk);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de fallos ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
plot(detectado_li);
hold on;
% plot(detectado_pbk);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de ciberataque ec.paridad tras ' num2str(ref_num_fallos) ' fallos '↙
'p= '  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');

% Segunda figura con tres subplots
figure(2);
subplot(3, 1, 1);
plot(beta_lik);
hold on;
plot(gamma_lik);
plot(r_Plikf);
xlim([4000 5500]);
```

```
title(['Ciberataque Litio Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
% plot(rb_pb);
% hold on;
plot(rb_lik);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de fallos Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
% plot(detectado_pb);
% hold on;
plot(detectado_lik);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de ciberataque Kalman tras ' num2str(ref_num_fallos) ' fallos ' 'p=↙
'  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');
```

```
clear;
close all;

load("residuos24h.mat"); load("Res_ec_paridad_ciber_sc.mat");
load("residuos24h_kalman.mat"); load("Res_ec_kalman_ciber_sc.mat");
r_Psc=r_nofault.data(:,4); %Ciberataque entre t=4500 y t=5000
r_Pscf=r.data(:,4);
r_Psck=r_kalmannofault.data(:,4);
r_Psckf=r_kalman.data(:,4);
p=0.75;

desviacion_sc=std(r_Psc);
desviacion_sck=std(r_Psck);


%calculo medias ventana (t+6)

for i=1:(length(r_Psc)-6)
    media_sc(i)=mean(r_Psc(i:(6+i)));
    media_sck(i)=mean(r_Psck(i:(6+i)));
end


for t=1:length(media_sc)

beta_sc(t)=norminv(p,media_sc(t), desviacion_sc);
gamma_sc(t)=norminv(1-p,media_sc(t), desviacion_sc);

beta_sck(t)=norminv(p,media_sck(t), desviacion_sck);
gamma_sck(t)=norminv(1-p,media_sck(t), desviacion_sck);


end

% for t=1:length(r_Pscf)
%     if r_Pscf(t)>beta_sc(t) | r_Pscf(t)<gamma_sc(t)
%         rb_sc(t)=1;
%     else
%         rb_sc(t)=0;
%     end
% end
% for t=1:length(r_Psckf)
%     if r_Psckf(t)>beta_sck(t) | r_Psckf(t)<gamma_sck(t)
%         rb_sck(t)=1;
%     else
%         rb_sck(t)=0;
%     end
% end


ref_num_fallos = 5;

cuenta_sc = 0;
for t=1:length(r_Pscf)
    if r_Pscf(t)>beta_sc(t) | r_Pscf(t)<gamma_sc(t)
```

```
        rb_sc(t)=1;
        if t==1
            rb_sc_ant=0;
        else
            rb_sc_ant=rb_sc(t-1);
        end
        if rb_sc(t)==1 && rb_sc_ant==0
            cuenta_sc = cuenta_sc + 1;
        end
    else
        rb_sc(t)=0;
    end
    if cuenta_sc >= ref_num_fallos
        detectado_sc(t) = 1;
    else
        detectado_sc(t) = 0;
    end
end

cuenta_sck = 0;
for t=1:length(r_Psckf)
    if r_Psckf(t)>beta_sck(t) | r_Psckf(t)<gamma_sck(t)
        rb_sck(t)=1;
        if t==1
            rb_sck_ant=0;
        else
            rb_sck_ant=rb_sck(t-1);
        end
        if rb_sck(t)==1 && rb_sck_ant==0
            cuenta_sck = cuenta_sck + 1;
        end
    else
        rb_sck(t)=0;
    end
    if cuenta_sck >= ref_num_fallos
        detectado_sck(t) = 1;
    else
        detectado_sck(t) = 0;
    end
end

for t = 4000:5500

dif_superior_paridad(t) = r_Pscf(t) - beta_sc(t) ;

end

for t= 4000:5500
umbral_up(t)   = -3*10^-4  ;
umbral_down(t) = -6*10^-4  ;

end
% figure; plot(beta_sc);title('Ciberataque Supercondensador ec.paridad p=0.65');↙
xlabel('Tiempo(segundos)');ylabel('Residuo'); hold on; plot(gamma_sc); plot(r_Pscf);↙
xlim([3620 6000]);
```

```
% figure; plot(beta_sck);title('Ciberataque Supercondensador Kalman p=0.65'); xlabel('Tiempo(segundos)');ylabel('Residuo');  hold on; plot(gamma_sck); plot(r_Psckf); xlim([3620 6000]);
% figure; plot(rb_sc);title('Detección de ciberataque ec.paridad p=0.65'); xlabel('Tiempo(segundos)');ylabel('Existencia de fallos'); xlim([3620 6000]); ylim([-0.5 1.5]);
% figure; plot(rb_sck);title('Detección de ciberataque Kalman p=0.65'); xlabel('Tiempo(segundos)');ylabel('Existencia de fallos'); xlim([4000 5500]); xlim([3620 6000]); ylim([-0.5 1.5]);

 % figure(1); plot(beta_sc); title(['Ciberataque supercondensador ec.paridad p= ' num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Residuo','FontSize', 20); hold on; plot(gamma_sc); plot(r_Pscf); xlim([4000 5500]);
% figure(2); plot(beta_sck); title(['Ciberataque supercondensador Kalman p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Residuo','FontSize', 20); hold on; plot(gamma_sck); plot(r_Psckf); xlim([4000 5500]);
% figure(3); plot(rb_sc); title(['Detección de fallos ec.paridad p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de fallos','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);
% figure(4); plot(rb_sck); title(['Detección de fallos Kalman p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de fallos','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);
% figure(5); plot(detectado_sc);title(['Detección de ciberataque ec.paridad tras ' num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de ciberataque','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);
% figure(6); plot(detectado_sck);title(['Detección de ciberataque Kalman tras ' num2str(ref_num_fallos) ' fallos ' 'p= '  num2str(p)],'FontSize', 25); xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Existencia de ciberataque','FontSize', 20); xlim([4000 5500]); ylim([-0.5 1.5]);


% figure(7); plot(dif_superior_paridad);title('Diferencia entre residuo y umbral superior');xlabel('Tiempo(segundos)','FontSize', 20);ylabel('Valor diferencia','FontSize', 20); hold on; plot(umbral_up); plot(umbral_down) ; xlim([4005 5500]);


% Primera figura con tres subplots
figure(1);
subplot(3, 1, 1);
plot(beta_sc);
hold on;
plot(gamma_sc);
plot(r_Pscf);
xlim([4000 5500]);
title(['Ciberataque Supercondensador ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
plot(rb_sc);
hold on;
% plot(rb_pbk);
```

```
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de fallos ec.paridad p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
plot(detectado_sc);
hold on;
% plot(detectado_pbk);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de ciberataque ec.paridad tras ' num2str(ref_num_fallos) ' fallos '↙
'p= '  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');

% Segunda figura con tres subplots
figure(2);
subplot(3, 1, 1);
plot(beta_sck);
hold on;
plot(gamma_sck);
plot(r_Psckf);
xlim([4000 5500]);
title(['Ciberataque Supercondensador Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Residuo');

subplot(3, 1, 2);
% plot(rb_pb);
% hold on;
plot(rb_sck);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de fallos Kalman p= ' num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de fallos');

subplot(3, 1, 3);
% plot(detectado_pb);
% hold on;
plot(detectado_sck);
xlim([4000 5500]);
ylim([-0.5 1.5]);
title(['Detección de ciberataque Kalman tras ' num2str(ref_num_fallos) ' fallos ' 'p=↙
'  num2str(p)]);
xlabel('Tiempo(segundos)');
ylabel('Existencia de ciberataque');
```