



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

CARACTERIZACIÓN ESTADÍSTICA DEL PERFIL DE USO DE BATERÍAS PARA EL
PRONÓSTICO DEL ESTADO-DE-CARGA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

HUGO FERNANDO NAVARRETE ECHEVERRÍA

PROFESOR GUÍA:
MARCOS ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:
JORGE SILVA SÁNCHEZ
HÉCTOR AGUSTO ALEGRÍA

SANTIAGO DE CHILE
JULIO 2014

RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO
POR: HUGO FERNANDO NAVARRETE ECHEVERRÍA
FECHA: JULIO 2014
PROF. GUÍA: MARCOS ORCHARD CONCHA

CARACTERIZACIÓN ESTADÍSTICA DEL PERFIL DE USO DE BATERÍAS PARA EL PRONÓSTICO DEL ESTADO-DE-CARGA

En los últimos años hubo una fuerte tendencia hacia la utilización de equipos cuya fuente de energía corresponde a baterías o dispositivos acumuladores de energía. Sin embargo, el uso de algunos de estos equipos, correspondientes por lo general a gran maquinaria, se ha visto limitada por la falta de métodos capaces de garantizar un cierto nivel en su autonomía. Un caso particular de este tipo de equipos corresponde a los automóviles eléctricos; cuya industria ha puesto principales esfuerzos en desarrollar sistemas de administración de baterías. Uno de los principales objetivos de estos sistemas es poder asegurar al usuario la duración de la energía disponible en el automóvil. Para ellos se hace necesario poder obtener información sobre la cantidad de energía remanente en su batería, parámetro conocido en la literatura como el Estado-de-Carga.

El conocimiento del parámetro anteriormente mencionado por sí solo no es suficiente, ya que para poder garantizar que el sistema posee una cierta autonomía es necesario obtener información sobre el comportamiento futuro del sistema. Este último problema se refiere al pronóstico de dicho estado. Diversos enfoques se encuentran disponibles en la literatura para su resolución; sin embargo, un desafío importante corresponde a la definición del perfil de consumo con que es utilizado el dispositivo durante el horizonte futuro.

Es en esta última problemática en que se enfoca el trabajo realizado. En particular se plantea una metodología que modela el perfil de uso de la batería mediante una cadena de Markov homogénea, definiendo los parámetros que la representan. Este modelo corresponde a una caracterización estadística de dicho perfil de uso, la cual es obtenida con los datos de corriente anteriores al instante de predicción. El objetivo de la realización de esta caracterización estadística es poder obtener la mayor información acerca de cómo ha sido descargada la batería, y utilizar dicha información para el perfil de uso durante la etapa de predicción.

Para validar la metodología se utilizaron dos conjuntos de datos correspondientes al proceso de descarga de baterías de Ion-Litio; además de un módulo de pronóstico del Estado-de-Carga basado en Filtro de Partículas. Los resultados obtenidos en la etapa de pronóstico fueron evaluados mediante medidas de desempeño adecuadas, mostrando que la caracterización estadística obtenida mediante la metodología presentada es de utilidad para la resolución de la problemática asociada a la predicción del Estado-de-Carga y estimación de la autonomía de equipos energizados con baterías.

Como principal conclusión de este trabajo se obtiene una novedosa metodología, sustentada en herramientas matemáticas, que es capaz caracterizar un perfil de descarga de una batería mediante una cadena de Markov homogénea. El desarrollo de esta metodología no es específico para estos perfiles, por lo que puede ser utilizada para la caracterización de una serie de tiempo en general.

A mis padres y a mi abuela María Luisa.

Agradecimientos

Primero que todo quiero agradecer a mi familia. Especialmente a mis padres, Fresia y Hugo, quienes siempre me han apoyado y ayudado. También a mi abuela María Luisa, con quien desde muy pequeño hemos tenido un fuerte lazo, siendo una segunda madre para mí. Muchas gracias también a Carmen, quien siempre me ha apoyado en todas las etapas de mi vida.

Agradezco al profesor Marcos Orchard, quien guió este trabajo, y sin su valiosa ayuda éste no habría llegado a un buen término. También agradecer toda la disposición que ha tenido conmigo, las oportunidades y confianza dadas durante mi paso por la universidad.

Quiero agradecer a María Paz, quien me acompañó y apoyó durante los últimos años. Muchas gracias por todo.

Agradecer también a los profesores Jorge Silva y Héctor Agosto por la contribución dada para realizar este trabajo.

Por último agradezco a mis amigos, tanto los conocidos en el paso por la universidad, los del colegio y los conocidos en la vida en general.

Tabla de contenido

1. Introducción	1
1.1. Objetivos Generales de este Trabajo	2
1.2. Objetivos Específicos de este Trabajo	2
2. Marco Teórico	4
2.1. Antecedentes	4
2.2. K-means	6
2.3. Estimador de Máxima Verosimilitud	7
2.4. Test de Hipótesis	9
2.4.1. Test de Bondad de Ajuste	11
2.4.2. Tablas de Contingencia	13
2.5. Desigualdades de Concentración	17
2.6. Cadenas de Markov	18
2.7. Inferencia Bayesiana y Métodos de Monte Carlo	20
2.7.1. Inferencia Bayesiana	20
2.7.2. Métodos Secuenciales de Monte Carlos: Filtro de Partículas	22
2.7.3. Esquema de Pronóstico Basado en Filtro de Partículas para Sistemas Dinámicos No-Lineales	25
3. Desarrollo Metodológico	29
3.1. Inferencia Estadística sobre Cadenas de Markov	29
3.1.1. Estimación de las Probabilidades de Transición	29
3.1.2. Test de Hipótesis de que la Cadena es de un Orden Determinado	30
3.2. Conjunto de Datos Utilizados	31
3.3. Elección de los Parámetros de la Cadena de Markov	32
3.3.1. Determinación del Valor de los Estados	33
3.3.2. Determinación de las Probabilidades de Transición	34
3.3.3. Determinación de la Cantidad de Estados	35
3.4. Elección del Orden de la Cadena de Markov	38
4. Validación de la Metodología	43
4.1. Modelo de Estado-Espacio para Estimación y Pronóstico del SOC Utilizando Caracterización Estadística de su Perfil de Uso Futuro	43
4.2. Medidas de Desempeño para Módulos de Pronóstico Basados en Filtro de Partículas	46
4.3. Determinación de los Parámetros del Algoritmo de Filtro de Partículas Utilizado . .	47

4.4. Resultados del Módulo de Pronóstico Basado en Filtro de Partículas y Caracterización Estadística del Perfil de Uso Futuro	48
5. Conclusiones	56
Bibliografía	58

Índice de tablas

4.1. Parámetros del modelo para las baterías asociadas a los conjuntos de datos Set #1 y Set #2.	45
4.2. Resultados de diferentes realizaciones del módulo de pronóstico para el SOC (Conjunto de datos Set #1). Ground Truth del EOD: 2738 [seg].	49
4.3. Resultados de diferentes realizaciones del módulo de pronóstico para el SOC (Conjunto de datos Set #2). Ground Truth del EOD: 3381 [seg].	50

Índice de figuras

3.1.	Conjunto de datos de voltaje y corriente correspondientes a pruebas experimentales de descarga de un ESD.	32
3.2.	Datos de Corriente Set #1.	33
3.3.	Acercamiento de Datos de Corriente Set #1, ilustrando dos situaciones irregulares.	34
3.4.	Preprocesamiento de la cadena de Markov correspondiente Datos de Corriente Set #1.	35
3.5.	Comportamiento de las cotas superiores en función del número de observaciones, utilizando $t = 0.1$	37
3.6.	Diagrama de flujo que grafica la metodología descrita.	39
3.7.	Perfil de corriente correspondientes al Set #1 y su trayectoria asociada a una cadena de Markov.	41
3.8.	Acercamiento del perfil de corriente correspondientes al Set #2 y su trayectoria asociada a una cadena de Markov.	42
4.1.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): SOC estimado (línea negra), SOC pronosticado (línea punteada magenta), e intervalo de confianza del 95 % para la predicción del SOC (línea punteada verde).	49
4.2.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): a) Evolución del estado x_1 , en función del SOC, durante la etapa de estimación (línea roja) y predicción (línea punteada magenta). b) Voltaje medido (línea azul), voltaje estimado (línea cyan), y voltaje predicho (línea punteada magenta).	50
4.3.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): Estimación de la pdf del EOD (normalizada). Líneas verticales punteadas indican los límites del intervalo de confianza del 95 % (líneas exteriores) y la esperanza (línea interior) para esta pdf.	51
4.4.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): SOC estimado (línea negra), SOC pronosticado (línea punteada magenta), e intervalo de confianza del 95 % para la predicción del SOC (línea punteada verde).	52
4.5.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): a) Evolución del estado x_1 , en función del SOC, durante la etapa de estimación (línea roja) y predicción (línea punteada magenta). b) Voltaje medido (línea azul), voltaje estimado (línea cyan), y voltaje predicho (línea punteada magenta).	53

4.6.	Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): Estimación de la pdf del EOD (normalizada). Líneas verticales punteadas indican los límites del intervalo de confianza del 95 % (líneas exteriores) y la esperanza (línea interior) para esta pdf.	54
4.7.	Perfil de corriente correspondiente al Set #1 y la trayectoria asociada a una cadena de Markov (etapa de estimación); junto con una realización de la cadena de Markov (etapa de pronóstico).	54
4.8.	Acercamiento del perfil de corriente correspondiente al Set #2 y la trayectoria asociada a una cadena de Markov (etapa de estimación); junto con una realización de la cadena de Markov (etapa de pronóstico).	55

Capítulo 1

Introducción

Durante los últimos años se ha puesto mucho interés en el estudio de los dispositivos almacenadores de energía (ESD, por sus siglas en inglés) dado que se ha producido una creciente tendencia hacia su utilización, tanto como fuente principal de energía o como fuente secundaria, en la mayoría de los equipos electrónicos. Los más comunes de estos equipos corresponden a cámaras digitales, *smartphones*, computadores portátiles, *tablets*, entre otros. Para este tipo de aplicaciones, disponer de un sistema supervisor de la batería no es de vital importancia. Una muestra de la poca importancia en la preocupación de disponer de información certera sobre la autonomía de la batería se manifiesta en lo poco sofisticado de los sistemas de monitoreo presentes en estos equipos. Esto se debe a que una información incorrecta sobre la autonomía del dispositivo se traduce, en general, sólo en que este se apague y haya que volver a cargarlo. Sin embargo, existen otros tipos de equipos o máquinas, en donde disponer de un monitoreo confiable de su sistema suministrador de energía es crítico. Por ejemplo, la obtención de información errónea sobre el estado de la batería utilizada por robots exploradores empleados en las misiones espaciales puede causar el completo incumplimiento de ésta. Otro caso crítico corresponde al de los vehículos aéreos no tripulados (UAV, por sus siglas en inglés), en donde una mala planificación de su uso puede causar la destrucción del equipo, causando grandes pérdidas monetarias.

Si bien las aplicaciones anteriormente mencionadas no son típicas en el ámbito civil, hay una de particular interés que ha tomado fuerza en los últimos años. Esta corresponde al monitoreo del banco de baterías presente en los autos eléctricos (EV, por sus siglas en inglés). Hoy en día la penetración de estos vehículos en el mercado mundial es aún baja. Sin embargo, con la creciente preocupación sobre el calentamiento global y el agotamiento de los combustibles fósiles se ha iniciado un interés en la investigación y desarrollo de este tipo de vehículo en desmedro de los que utilizan un motor de combustión interna [1].

Aunque este cambio en el uso del tipo de vehículo sucederá inevitablemente algún día, existen importantes desafíos que deben ser superados para garantizar su éxito. Uno de estos desafíos corresponde al problema de “*range anxiety*”, el cual se refiere al temor del usuario a quedarse con el auto desenergizado en el camino [2]. Este temor se incrementa dado el hecho de que actualmente las estaciones de servicio que disponen de cargadores eléctricos son poco comunes. Es por esto que

se ha puesto especial interés en el eficiente desarrollo de sistemas de administración de baterías (BMS, por sus siglas en inglés) [3]. Estos sistemas tienen como principal objetivo: *(i)* proporcionar información en tiempo real; *(ii)* reducir el periodo de carga de la batería, garantizando que esta acumule la suficiente energía para garantizar la autonomía del sistema; *(iii)* maximizar el número total de ciclos durante los cuales puede ser utilizada la batería; y *(iv)* maximizar el tiempo de uso de la batería durante un ciclo particular de operación. Para lograr esto, los BMS deben considerar información acerca de parámetros de la batería; tales como el Estado-de-Carga (SOC, por sus siglas en inglés) [4], [3], el Estado-de-Salud (SOH, por sus siglas en inglés) [5], [3], [6], y su vida útil remanente (RUL, por sus siglas en inglés) [7].

El Estado-de-Carga corresponde a un indicador que representa la cantidad de energía disponible en la batería, con respecto a su máxima energía que es capaz de almacenar. Información acerca de esta variable es fundamental para poder planificar de forma óptima la utilización de un sistema energizado mediante un ESD en general; y particularmente en el caso de automóviles eléctricos.

1.1. Objetivos Generales de este Trabajo

A pesar de que un conocimiento certero del SOC corresponde a una información muy útil, no es suficiente para poder realizar decisiones con respecto a la autonomía del sistema. Para esto es necesario además tener conocimiento sobre el comportamiento futuro de este parámetro. Esta problemática se refiere al pronóstico del SOC. Sin embargo, para poder desarrollarla es necesario conocer el perfil de uso que se le dará al ESD en dicha etapa, lo que en la mayoría de las aplicaciones no es posible. Es por esto que el objetivo general del presente trabajo corresponde al desarrollo de una metodología que permita obtener una caracterización estadística del perfil de uso de la batería durante la etapa previa a la realización del pronóstico. Esta caracterización, asumiendo que representa de forma adecuada a la utilización futura de la batería, es empleada en la etapa de pronóstico para simular el perfil al que se somete el dispositivo. De esta forma es posible estimar el comportamiento que tendrá dicho parámetro, y poder así evaluar la factibilidad de llevar a cabo un determinado uso para el dispositivo.

1.2. Objetivos Específicos de este Trabajo

Los objetivos específicos de este trabajo corresponden al diseño de la metodología anteriormente señalada, detallando cada uno de los pasos y necesarios para llevarla a cabo. El enfoque que se adopta corresponde a realizar una caracterización estadística del perfil de corriente de descarga de un ESD mediante una cadena de Markov homogénea. La metodología debe entregar los parámetros de dicha cadena que mejor representan al perfil de corriente. Se requiere que el enfoque con que se hace la elección de dichos parámetros sea justificado mediante argumentos basados en herramientas matemáticas, con el objetivo de garantizar un cierto grado de formalidad.

Una vez diseñada la metodología, para su validación se utilizará dentro de un módulo de pro-

nóstico basado en Filtro de Partículas, junto con datos correspondientes a pruebas experimentales del proceso de descarga de baterías. De esta forma se podrá obtener resultados para el pronóstico del SOC; y evaluar, mediante medidas de desempeño adecuadas, si efectivamente esta metodología es útil para realizar una caracterización estadística del perfil de uso de baterías enfocada en el pronóstico del SOC.

La implementación de esta metodología, junto con la del módulo de pronóstico, se desarrolla mediante el *software* computacional MATLAB®.

La estructura de este trabajo es la siguiente: En el Capítulo 2 se presentan los antecedentes sobre la problemática de estimación y pronóstico del SOC. También se describen las herramientas matemáticas que son necesarias para llevar a cabo el desarrollo, y su posterior validación, de la metodología presentada. En el Capítulo 3 se describe en detalle la metodología, junto con dar argumentos de las consideraciones realizadas. En el Capítulo 4 se realiza la validación de la metodología, a través de su incorporación en un módulo de pronóstico para el SOC. Se explican los detalles de la implementación de dicho módulo y los resultados objetivos, junto con presentar medidas de desempeño para su evaluación. Finalmente, en el Capítulo 5 se presentan las conclusiones obtenidas a través de este trabajo, junto con plantear líneas a seguir para la realización de trabajos futuros.

Capítulo 2

Marco Teórico

2.1. Antecedentes

En la actualidad, el estudio de parámetros críticos de los ESD, tales como el Estado-de-Carga presenta tres importantes aspectos que deben ser considerados para obtener un eficiente enfoque de pronóstico: (i) cómo modelar el ESD, (ii) cómo realizar la estimación en tiempo real del SOC, y (iii) cómo predecir el SOC. En este último punto es de particular interés estudiar el impacto que tienen los perfiles de descarga futuros con el objetivo de obtener estimaciones confiables para el punto de descarga (EOD, por sus siglas en inglés). Si bien estos tres aspectos son problemas; hoy en día, abiertos, los dos primeros han sido abordados mediante diversos enfoques. El tercer aspecto mencionado corresponde a la principal problemática a tratar en este trabajo, ya que en la etapa de predicción no se conoce con certeza el perfil de utilización al cual es sometido el ESD, requiriendo una caracterización estadística del mismo para incorporar dicha información para realizar un eficiente pronóstico del tiempo de autonomía del dispositivo que utiliza el ESD.

Una de las principales dificultades para estimar el SOC es que éste no es medible u observable directamente; debiendo ser inferido por métodos indirectos, los que utilizan información de otros parámetros de la batería (corriente de descarga, voltaje en bornes, temperatura, entre otros [3], [8], [9], [10], [11], [12]). Por otro lado, es necesario disponer de un modelo que caracterice el comportamiento de la batería; tales como modelos empíricos, estocásticos, y fenomenológicos/electroquímicos. Estos últimos requiere de mediciones muy precisas y utilización de un gran número de parámetros, implicando altos costos computacionales y limitando su uso a aplicaciones de tiempo real [3], [13]. También existen modelos para el voltaje en circuito abierto (OCV, por sus siglas en inglés) [11]. Si bien su aplicación es sencilla, requiere de largos tiempos de reposo de la batería, dificultando su aplicación cuando no se cuenta con dicho recurso [3], [13], [10], [11], [12]. Por su parte, los modelos basados en espectroscopía de impedancia electroquímica han sido bastante exitosos [3], [14], [12]; sin embargo, requieren de costosos equipos, limitando su masificación [15].

En la actualidad existe un especial interés por los modelos empíricos, debido a su flexibilidad y posibilidad de trabajar con datos limitados y ruidosos. Dentro de estos modelos se encuentran los

basados en la respuesta al impulso [4], lógica difusa [6], redes neuronales [3], [13], [9], [16], técnicas basadas en sistemas de parámetros linealmente variables [17], y máquinas de soporte vectorial [18].

Dentro del contexto de los modelos fenomenológicos se encuentra el enfoque de modelar el ESD mediante un circuito eléctrico equivalente, para caracterizar el comportamiento fisicoquímico presente en ellos [5]. Así, se obtiene un modelo no lineal para representar el voltaje en bornes y el SOC de la batería, el cual sólo requiere de mediciones de ciertos parámetros presentes en la batería; tales como su corriente de descarga y voltaje, posibilitando su aplicación en línea.

Dada la no linealidad de los modelos utilizados por este último enfoque, uno de los métodos utilizados en la literatura, tanto para realizar propósitos de estimación y pronóstico, corresponden al Filtro Extendido de Kalman (EKF, por sus siglas en inglés) [19]. Sin embargo, este método presenta la desventaja de producir errores, debido a la linealización del sistema, al realizar la etapa pronóstico en horizontes de predicción largos. Una alternativa que ha probado ser eficiente para realizar estimación y pronóstico en modelos no lineales corresponde a los métodos secuenciales de Monte Carlo (SMC, por sus siglas en inglés) [5], [7], [10], [15], [20], [21]; también conocido como Filtro de Partículas. Una de las ventajas que posee estos métodos, es su capacidad para adaptar parámetros del modelo durante la etapa de estimación, lo que es crítico para poder realizar una adecuada etapa de pronóstico.

La gran mayoría de los enfoques de pronóstico para el SOC consideran un perfil de consumo futuro constante, el cual corresponde típicamente al valor promedio de los niveles de corriente medidos en los instantes previos al tiempo de pronóstico. Esta metodología corresponde a un supuesto débil y descartable para ser aplicado en equipos alimentados por baterías, los cuales generalmente requieren una alimentación de potencia constante, ya que a medida que el acumulador se descarga durante su utilización, el voltaje en sus bornes decae, requiriendo un mayor consumo de corriente, traduciéndose en realizar la etapa de pronóstico con poca exactitud [22].

Por otro lado, en [23] se ha estudiado un enfoque de pronóstico para el SOC, basado en algoritmos entrenados con datos de múltiples descargas a diferentes niveles de corriente constante. Este enfoque utiliza un perfil de descarga variante en el tiempo para la etapa de predicción. Sin embargo, se asume que dicho perfil de descarga es determinístico y conocido al momento de predicción, lo cual en muchas aplicaciones de interés no es posible conocer.

Dentro de los enfoques que se han implementado para incluir una caracterización estadística del perfil de utilización futuro para la etapa de pronósticos, puede mencionarse la definición de un rango de operación para los valores máximos y mínimos del perfil de descarga utilizado por el ESD, junto con la información de los datos correspondiente a dicho perfil. Lo anterior permite la generación de una distribución empírica de los valores que caracterizan la incertidumbre del perfil de descarga futuro de la batería [24], [22].

Para desarrollar el diseño de la metodología propuesta en este trabajo, la cual permite obtener

una caracterización estadística del perfil de descarga de un ESD, es necesario la utilización de herramientas matemáticas. Estas herramientas son descritas a continuación.

2.2. K-means

Una de los problemas más importantes en el área de minería de datos corresponde a realizar un agrupamiento de ellos, también conocido como *clustering*. Este problema tiene por objetivo realizar la partición de un conjunto de n datos multidimensionales en k grupos; en donde cada dato pertenece al grupo cuya distancia entre su media y el dato sea menor. El número de grupos corresponde a un parámetro del problema, el cual debe ser definido de antemano [25].

Sea $D = \{x_1, \dots, x_n\}$ el conjunto de datos; y $C = \{C_1, \dots, C_k\}$ una colección de conjuntos, que corresponden a una partición del conjunto de datos. Se define al función de error como:

$$E = \sum_{i=1}^k \sum_{x \in C_i} d\{x, \mu(C_i)\}. \quad (2.1)$$

En donde, $\mu(C_i)$ corresponde a la media, o centroide, del conjunto C_i ; y $d\{x, \mu(C_i)\}$ es la distancia entre los puntos x y $\mu(C_i)$. En aplicaciones estándar, la distancia utilizada por defecto corresponde a la distancia Euclidiana.

Matemáticamente, el problema de agrupamiento puede ser formulado mediante el siguiente problema de optimización:

$$\begin{aligned} & \underset{W, Q}{\text{minimize}} && \sum_{i=1}^n \sum_{l=1}^k w_{il} \cdot d\{x_i, q_l\} \\ & \text{subject to} && \sum_{l=1}^k w_{il} = 1, \quad (\forall i \in \{1, \dots, n\}). \\ & && w_{il} \in \{0, 1\}, \quad (\forall i \in \{1, \dots, n\}, \forall l \in \{1, \dots, k\}). \end{aligned}$$

En donde, $Q = \{q_1, \dots, q_k\}$, y $W[i, l] = w_{i,l}$.

Uno de los algoritmos más utilizados para resolver este problema; dado el buen desempeño que presenta en aplicaciones reales, a pesar de entregar una solución subóptima de éste, corresponde al algoritmo *K-means* [26]. Las principales ventajas de este algoritmo corresponde a su facilidad de ser programado, junto con una económica utilización de recursos computacionales; y por lo mismo, permite trabajar con grandes conjuntos de datos. Algunas de sus aplicaciones más utilizadas son: predicción no lineal, aproximación de distribuciones multivariadas, agrupamiento por similitud, entre otras.

K-means se basa en, dado un valor para k cualquiera, asociar iterativamente los elementos del conjunto de datos al centroide más cercano, los cuales son recalculados en cada iteración del algoritmo. Este proceso se repite hasta que la función de error (2.1) no presente variación.

2.3. Estimador de Máxima Verosimilitud

En estadística, un problema frecuente corresponde a la extracción de características propias de un fenómeno aleatorio en base a observaciones de variables aleatorias generadas por el mismo. Se define por modelo estadístico como la identificación de estas variables aleatorias de interés para el proceso, la especificación de la distribución conjunta o una familia de posibles distribuciones conjuntas para dichas variables aleatorias, la identificación de los parámetros desconocidos para dichas distribuciones, y la especificación de la distribución de tales parámetros desconocidos. Además, se define por inferencia estadística [27] al procedimiento en donde se plantea una afirmación probabilística sobre algún modelo estadístico. Por afirmación probabilística se entiende una afirmación en donde se hace uso de algún concepto referente a la teoría de probabilidades.

En un problema de inferencia estadística, un parámetro corresponde a una característica (o combinación de ellas) que determina la distribución conjunta de las variables aleatorias de interés. El conjunto de todos los posibles valores para un parámetro, escalar o vectorial, se conoce como espacio de parámetros y se denota por Ω .

Sea X_1, \dots, X_n las observaciones de una variable de interés en un problema de inferencia estadística, y sea r una función a valores reales de n variables reales. Se define como un estadístico a la variable aleatoria $T = r(X_1, \dots, X_n)$. De esta forma, un estadístico resume la información disponible en las observaciones de la variable aleatoria; y es por esto que muchos de los métodos utilizados para realizar inferencia estadística se basan en el uso de estadísticos adecuados.

Uno de los métodos más utilizados en inferencia estadística para la estimación de parámetros corresponde al método de estimación de máxima verosimilitud [27], [28]. Esto se debe a que, junto con entregar resultados razonables para el problema de estimación, su obtención es simple e intuitiva.

Sea X_1, \dots, X_n muestras aleatorias de una distribución de probabilidad, discreta continua, cuya función de probabilidad o función de densidad de probabilidad es $f(x|\theta)$, en donde el parámetro θ , escalar o vectorial, pertenece a un espacio de parámetros Ω . Para cada vector de observaciones $\mathbf{x} = (x_1, \dots, x_n)$, la función probabilidad conjunta o función de densidad de probabilidad conjunta corresponde a $f_n(\mathbf{x}|\theta)$.

Cuando la función $f_n(\mathbf{x}|\theta)$ de las observaciones en una muestra aleatoria es considerada función

del parámetro θ , para unos valores de x_1, \dots, x_n dados, dicha función es conocida como la función de verosimilitud.

Para cada posible valor del vector de observaciones \mathbf{x} , se denota por $\delta(\mathbf{x}) \in \Omega$ al valor de $\theta \in \Omega$ que maximiza la función de verosimilitud, y sea $\hat{\theta} = \delta(\mathbf{X})$ un estimador para tal θ . Este estimador $\hat{\theta}$ es conocido como un estimador de máxima verosimilitud de θ . Una vez que $\mathbf{X} = \mathbf{x}$ es observado, el valor $\delta(\mathbf{x})$ es llamado una estimación de máxima verosimilitud de θ .

En algunos problemas, para ciertos vectores de observaciones \mathbf{x} , el máximo valor de $f_n(\mathbf{x}|\theta)$ puede no ser alcanzado por algún valor de $\theta \in \Omega$. En estos casos, no existe un estimador de máxima verosimilitud para θ . Por otro lado, para otros valores observados de \mathbf{x} , el máximo valor de $f_n(\mathbf{x}|\theta)$ puede ser alcanzado por más de un valor del espacio de parámetros Ω . En tales casos, el estimador de máxima verosimilitud no es único, y cualquiera de tales valores puede ser elegido como el valor del estimador $\hat{\theta}$. Sin embargo, en muchos de los problemas el estimador de máxima verosimilitud existe y es único.

Bajo supuestos débiles, los cuales son típicamente satisfechos en problemas prácticos, las principales propiedades que cumple este estimador son las siguientes [27], [29]:

- Consistencia: si el número de observaciones, n , con los que se calcula el estimador de máxima verosimilitud del parámetro θ se hace suficientemente grande; la secuencia formada por dichos estimadores, $\hat{\theta}_n$, corresponde a una secuencia consistente de estimadores de θ . Esto quiere decir que la secuencia $\hat{\theta}_n$ converge en probabilidad al verdadero parámetro θ .
- Invarianza: si $\hat{\theta}_n$ corresponde al estimador de máxima verosimilitud de θ , y si $g(\cdot)$ es una función inyectiva; entonces $g(\hat{\theta}_n)$ corresponde al estimador de máxima verosimilitud de $g(\theta)$.
- Normalidad asintótica: si $\hat{\theta}_n$ corresponde al estimador de máxima verosimilitud de θ obtenido mediante n observaciones; entonces, este estimador tiene la siguiente distribución asintótica:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, I^{-1}(\theta)). \quad (2.2)$$

En donde, $\mathcal{N}(\cdot, \cdot)$ corresponde a una distribución normal, y $I(\theta)$ es la matriz de información de Fisher:

$$I(\theta) = \mathbb{E} \left\{ \left[\frac{\partial}{\partial \theta} \log\{f(x|\theta)\} \right] \left[\frac{\partial}{\partial \theta} \log\{f(x|\theta)\} \right]' \right\}. \quad (2.3)$$

- Eficiencia asintótica: si $\hat{\theta}_n$ corresponde al estimador de máxima verosimilitud de θ obtenido mediante n observaciones; entonces este estimador es asintóticamente eficiente. Esto quiere decir que asintóticamente alcanza la cota inferior de Cramer-Rao al hacer tender el número de observaciones a infinito.

2.4. Test de Hipótesis

Un test de hipótesis corresponde a un procedimiento estadístico, en donde se determina si cierta propiedad supuesta para una población es compatible con lo observado en una muestra de ella. En específico, se tiene un parámetro desconocido θ perteneciente al espacio de parámetros Ω . Este espacio de parámetros puede ser particionado en dos subconjuntos disjuntos Ω_0 y Ω_1 , tal que $\Omega_0 \cap \Omega_1 = \emptyset$ y $\Omega_0 \cup \Omega_1 = \Omega$. La problemática a resolver mediante test de hipótesis corresponde a determinar si $\theta \in \Omega_0$ ó $\theta \in \Omega_1$.

Se considera H_0 la hipótesis de que $\theta \in \Omega_0$ y H_1 la hipótesis de que $\theta \in \Omega_1$. Una consecuencia inmediata de que los subconjuntos Ω_0 y Ω_1 sean disjuntos corresponde a que sólo una de las hipótesis H_0 ó H_1 sea verdadera. Un problema de este tipo, en donde hay que decidir entre dos posibles elecciones, es conocido como un problema de test de hipótesis [27], [30], [31]. Mientras que el procedimiento para decidir cual hipótesis elegir es conocido como un *test*.

La hipótesis H_0 es llamada hipótesis nula, mientras que la hipótesis H_1 es llamada hipótesis alternativa. Al realizar un test, si la decisión corresponde a $\theta \in \Omega_1$, se dice que se rechaza H_0 . Mientras que si la decisión es $\theta \in \Omega_0$, se dice que no se rechaza H_0 .

Si Ω_i , con $i = \{1, 0\}$, contiene sólo un valor para el parámetro θ , entonces se dice que H_i es una hipótesis simple. Si Ω_i contiene más de un valor para θ , entonces se dice que H_i es una hipótesis compuesta. Bajo una hipótesis simple la distribución de las observaciones es completamente determinada. Mientras que bajo una hipótesis compuesta sólo está determinado que dicha distribución pertenece a cierta clase.

En general, se formula el siguiente problema para la elección de la hipótesis:

$$H_0 : \theta \in \Omega_0, \quad \text{ó} \quad H_1 : \theta \in \Omega_1. \quad (2.4)$$

Para determinar qué elección hacer, se dispone de un vector $\mathbf{X} = (X_1, \dots, X_n)$ conformado por n muestras aleatorias correspondientes a observaciones de una distribución que depende del parámetro desconocido θ . Se denota por S al espacio muestral del vector aleatorio n -dimensional \mathbf{X} . Algunos test pueden consistir en realizar una partición de S en dos subconjuntos. Uno de estos subconjuntos S_1 contiene los valores de \mathbf{X} para los cuales se rechaza H_0 , y el otro subconjunto S_0 contiene los valores de \mathbf{X} para los cuales no se rechaza H_0 . Este subconjunto S_1 corresponde a la región crítica del test. De esta forma un test queda determinado por la especificación de su región crítica.

En muchos problemas relacionados con test de hipótesis la región crítica del test es definida mediante un estadístico, $T = r(\mathbf{X})$. Sea R un subconjunto de la recta de los reales. Si un test para (2.4) corresponde a: rechazar H_0 si $T \in R$, entonces T corresponde a un test estadístico, y R corresponde a la región de rechazo de dicho test. De esta forma cuando un test es definido mediante

un test estadístico T , y su correspondiente región de rechazo R , el conjunto $S_1 = \{x : r(x) \in R\}$ corresponde a su región crítica. Una vez definido el test estadístico a utilizar, es más simple trabajar en función de dicho test que desarrollar su región crítica.

Sea δ un test cualquiera. Se define $\pi(\theta|\delta)$ para cada valor de $\theta \in \Omega$ como la probabilidad de que el test δ rechace H_0 . La función $\pi(\theta|\delta)$ es conocida como la función de potencia del test δ . Una función potencia ideal es aquella en donde $\pi(\theta|\delta) = 0$ para cada valor de $\theta \in \Omega_0$, y $\pi(\theta|\delta) = 1$ para cada valor de $\theta \in \Omega_1$. De esta forma, independiente del valor real que posea el parámetro θ , siempre se decidirá la elección correcta con probabilidad 1.

Sin embargo, un test que posea la función potencia ideal anteriormente descrita rara vez existe. Esto hace que la probabilidad de cometer algún tipo de error sea no nula. Para cada valor de $\theta \in \Omega_0$, rechazar H_0 es una decisión incorrecta, lo cual se conoce como error tipo I o error del primer tipo. Similarmente, para cada valor de $\theta \in \Omega_1$, no rechazar H_0 es una decisión incorrecta, lo cual se conoce como error tipo II o error del segundo tipo.

Un test adecuado es aquel que posee una baja probabilidad de cometer error. Lo deseable es que la función potencia del test sea pequeña para valores de $\theta \in \Omega_0$, y grande para $\theta \in \Omega_1$. Sin embargo, estos dos requerimientos trabajan uno en contra del otro, por lo que es necesario llegar a un balance entre ambos. Uno de los métodos más utilizados para encontrar dicho balance entre ambos requerimientos corresponde elegir un número $\alpha_0 \in [0, 1]$ tal que:

$$\pi(\theta|\delta) \leq \alpha_0. \quad (2.5)$$

Entre los test que satisfacen (2.5), la elección de un test adecuado corresponde a aquel cuya función potencia es lo más grande posible para $\theta \in \Omega_1$.

Un test que satisface (2.5) es conocido como un test de nivel α_0 , y se dice que tiene un nivel de significancia α_0 . Mientras que el tamaño del test $\alpha(\delta)$ es definido como:

$$\alpha(\delta) = \sup_{\theta \in \Omega_0} \pi(\theta|\delta). \quad (2.6)$$

En caso de tener un problema de test de hipótesis como el descrito en (2.4) y un test estadístico T , el cual rechaza la hipótesis nula si $T \geq c$ para alguna constante c , la función potencia del test viene dada por $\pi(\theta|\delta) = \mathbb{P}\{T \geq c|\theta\}$. Si el test se diseña para tener un nivel de significancia α_0 se debe cumplir:

$$\sup_{\theta \in \Omega_0} \mathbb{P}\{T \geq c|\theta\} \leq \alpha_0. \quad (2.7)$$

Esta función potencia corresponde a una función no creciente de c , por lo que el lado izquierdo de (2.7) es satisfecho para grande valores de c , pero no para pequeños. En caso de requerir una

función potencia grande para $\theta \in \Omega_1$ se debe elegir el valor de c más pequeño que cumpla con (2.7).

En muchas aplicaciones es de interés especificar, junto al valor que toma el test estadístico para los datos disponibles, todos los valores de α_0 tales que se rechace la hipótesis H_0 con un nivel α_0 con dicho datos. El menor valor de estos niveles α_0 es conocido como el valor p .

2.4.1. Test de Bondad de Ajuste

Al utilizar los test de hipótesis más clásicos se asume que los datos observados provienen de una cierta familia de distribuciones de probabilidad cuya forma analítica es conocida, aunque algunos de sus parámetros pueden ser desconocidos. Otro tipo de problema importante en estadística corresponde a aquellos en donde no se asume que las observaciones provienen de una familia de distribuciones paramétricas en particular, sino que el interés se enfoca en realizar inferencias sobre la distribución de donde provienen dichas observaciones. Estos problemas son conocidos como problemas no paramétricos, y los métodos estadísticos utilizados corresponden a métodos no paramétricos.

Uno de los métodos no paramétricos más utilizados corresponde al test de bondad de ajuste [27], [30]. Este test se utiliza cuando se tiene la creencia de que los datos observados provienen de una distribución en específico, y se desea testear la hipótesis nula de que los datos efectivamente provienen de dicha distribución en contra de la hipótesis alternativa. Está basado en la calidad del ajuste que se obtiene entre la frecuencia de observaciones presentes en la muestra tomada y la frecuencia esperada obtenida de la distribución hipotética. Para este tipo de problemas se considera que cada dato observado puede ser clasificado como miembro de una clase o tipo, cuyo número es finito. Estas observaciones son conocidas como datos categóricos.

El Test

Si una población consiste en items de k tipos diferentes, y sea p_i corresponde a la probabilidad de que un item seleccionado al azar sea de tipo i ($i = 1, \dots, k$). Sean p_1^0, \dots, p_k^0 números tales que $p_i^0 \geq 0$ para $i = 1, \dots, k$ y $\sum_{i=1}^k p_i^0 = 1$, y se desea testear las siguientes hipótesis:

$$H_0 : p_i = p_i^0 \text{ para } i = 1, \dots, k. \quad (2.8)$$

$$H_1 : p_i \neq p_i^0 \text{ para al menos un valor de } i. \quad (2.9)$$

Se toma una muestra aleatoria de tamaño n de esta población, correspondiente a n observaciones independientes cuya probabilidad de pertenecer ser del tipo i ($i = 1, \dots, k$) es p_i . En base a dichas observaciones se desea testear las hipótesis (2.8) y (2.9).

Sea N_i el número de observaciones, dentro de la muestra aleatoria, que son del tipo i . Con esto, N_1, \dots, N_k corresponden a números enteros no negativos tales que $\sum_{i=1}^k N_i = n$, y (N_1, \dots, N_k) tiene una distribución multinomial de parámetros n y $\mathbf{p} = (p_1, \dots, p_k)$. Cuando se cumple la hipótesis nula, el número esperado de observaciones del tipo i es np_i^0 ($i = 1, \dots, k$). El siguiente estadístico, conocido como estadístico χ^2 :

$$Q = \sum_{i=1}^k \frac{(N_i - np_i^0)^2}{np_i^0}, \quad (2.10)$$

tiene la propiedad de que si se cumple H_0 y $n \rightarrow \infty$, entonces converge en distribución a una χ^2 con $k - 1$ grados de libertad. Bajo el supuesto de que la hipótesis nula es verdadera, la diferencia entre el número de observaciones obtenidas N_i y el número esperado de observaciones debería ser pequeño. Esto lleva a que se rechace H_0 cuando $Q \geq c$, en donde c corresponde a una constante apropiada. Si se requiere utilizar el test con un nivel de significancia α_0 , c corresponde al cuantil $1 - \alpha_0$ de la distribución χ^2 con $k - 1$ grados de libertad.

Este test puede ser adaptado para el caso de hipótesis compuestas, en donde la hipótesis nula corresponde a que los datos observados provienen de una familia paramétrica de distribuciones. La hipótesis alternativa en este caso es que los datos provienen de una distribución que no pertenece a dicha familia.

Nuevamente se considera una población que contiene ítems de k tipos diferentes, y que p_i denota la probabilidad de que un ítem seleccionado al azar sea del tipo i ($i = 1, \dots, k$). Sea N_1, \dots, N_k definido de igual forma que anteriormente. En vez de testear la hipótesis nula de que p_1, \dots, p_k tienen un valor específico, interesa testear la hipótesis compuesta de que p_1, \dots, p_k pertenecen a un cierto subconjunto específico de posibles valores. En particular, el interés corresponde a la problemática en donde la hipótesis nula es que los valores de p_1, \dots, p_k pueden ser representados en función de un número menor de parámetros.

Interesa testear la hipótesis de que cada probabilidad p_i ($i = 1, \dots, k$) puede ser representada por una función $\pi_i(\boldsymbol{\theta})$ de un vector de parámetros $\boldsymbol{\theta} = (\theta_1, \dots, \theta_s)$, en donde se asume que $s < k - 1$, y ninguna componente del vector $\boldsymbol{\theta}$ puede ser expresada en función de las otras $s - 1$. Sea Ω el espacio de parámetros s -dimensional de los posibles valores de $\boldsymbol{\theta}$. Se asume que las funciones $\pi_1(\boldsymbol{\theta}), \dots, \pi_k(\boldsymbol{\theta})$ siempre forman un conjunto factible de valores para p_1, \dots, p_k en el sentido de que para cada vector $\boldsymbol{\theta} \in \Omega$, $\pi_i(\boldsymbol{\theta}) > 0$ ($i = 1, \dots, k$) y $\sum_{i=1}^k \pi_i(\boldsymbol{\theta}) = 1$. De esta forma, las hipótesis a ser testeadas pueden ser escritas como:

$$H_0 \quad : \quad \text{Existe un valor de } \boldsymbol{\theta} \in \Omega \text{ tal que } p_i = \pi_i(\boldsymbol{\theta}) \quad (i = 1, \dots, k). \quad (2.11)$$

$$H_1 \quad : \quad \text{no se cumple } H_0. \quad (2.12)$$

Para desarrollar el test de bondad de ajuste para las hipótesis (2.11) y (2.12), el estadístico Q definido en (2.10) debe ser modificado; ya que np_i^0 , el número esperado de observaciones del tipo i , dentro de una muestra aleatoria de n observaciones no está completamente especificado por H_0 .

Se debe hacer la modificación de reemplazar np_i^0 por el estimador de máxima verosimilitud de este valor esperado asumiendo que la hipótesis nula es verdadera. Si $\hat{\theta}$ corresponde al estimador de máxima verosimilitud de θ basado en las observaciones N_1, \dots, N_k , entonces el siguiente estadístico:

$$Q = \sum_{i=1}^k \frac{(N_i - n\pi_i(\hat{\theta}))^2}{n\pi_i(\hat{\theta})} \quad (2.13)$$

tiene la propiedad de que cuando H_0 en (2.11) es verdadera y $n \rightarrow \infty$, entonces converge en distribución a una χ^2 con $k - 1 - s$ grados de libertad.

2.4.2. Tablas de Contingencia

En los casos en que las observaciones corresponden a un par de variables aleatorias (conformadas por dos componentes), cada una de estas observaciones puede ser clasificada de dos formas. Mediante una tabla, conocida como tablas de contingencia [27], es posible presentar las frecuencias observadas de las componentes de cada observación. En general, se consideran tablas conformadas por R filas y C columnas. Para $i = 1, \dots, R$ y $j = 1, \dots, C$, p_{ij} representa la probabilidad de que una observación aleatoria perteneciente a una población en estudio sea clasificada en la i -ésima fila y en la j -ésima columna de la tabla. Se denota por $p_{i\bullet}$ la probabilidad marginal de que una observación sea clasificada en la i -ésima fila de la tabla. Similarmente, se denota por $p_{\bullet j}$ la probabilidad marginal de que una observación sea clasificada en la j -ésima columna de la tabla. Se cumple que:

$$p_{i\bullet} = \sum_{j=1}^C p_{ij} \quad y \quad p_{\bullet j} = \sum_{i=1}^R p_{ij}.$$

Como la suma de las probabilidades de todas las celdas de la tabla de contingencia deben sumar 1, se debe cumplir que:

$$\sum_{i=1}^R \sum_{j=1}^C p_{ij} = \sum_{i=1}^R p_{i\bullet} = \sum_{j=1}^C p_{\bullet j} = 1.$$

Test de Independencia

El test de bondad de ajuste descrito puede ser utilizado para probar la hipótesis de que las dos variables de clasificación con las que se construye una tabla de contingencia son variables aleatorias independientes [27], [30].

Se considera que una muestra de n observaciones es tomada de una población. Sea N_{ij} ($i = 1, \dots, R$ y $j = 1, \dots, C$) el número de observaciones que son clasificadas en la i -ésima fila y j -ésima

columna de una tabla de contingencia. Se denota por $N_{i\bullet}$ al número total de observaciones que son clasificadas en la i -ésima fila. Similarmente, se denota por $N_{\bullet j}$ al número total de observaciones que son clasificadas en la j -ésima columna. Con esto se tiene que:

$$N_{i\bullet} = \sum_{j=1}^C N_{ij} \quad y \quad N_{\bullet j} = \sum_{i=1}^R N_{ij}. \quad (2.14)$$

$$\sum_{i=1}^R \sum_{j=1}^C N_{ij} = \sum_{i=1}^R N_{i\bullet} = \sum_{j=1}^C N_{\bullet j} = n. \quad (2.15)$$

La propiedad de independencia corresponde a que para cada valor de i ($i = 1, \dots, R$) y j ($j = 1, \dots, C$), la probabilidad de que una observación sea clasificada en la celda (i, j) de la tabla de contingencia es igual a la probabilidad de que dicha observación sea clasificada en la fila i por la probabilidad de que sea clasificada en la columna j . Es decir, en base a un conjunto de valores observados, se desea testear las siguientes hipótesis:

$$H_0 : p_{ij} = p_{i\bullet} p_{\bullet j} \text{ para } i = 1, \dots, R \text{ y } j = 1, \dots, C. \quad (2.16)$$

$$H_1 : \text{no se cumple } H_0. \quad (2.17)$$

Cada uno de los valores observados pertenece a una de las celdas de la tabla de contingencia. Cuando la hipótesis nula es verdadera, los valores desconocidos de p_{ij} han sido expresados en función de los parámetros desconocidos $p_{i\bullet}$ y $p_{\bullet j}$. Como $\sum_{i=1}^R p_{i\bullet} = 1$ y $\sum_{j=1}^C p_{\bullet j} = 1$, el número de parámetros que deben ser estimados cuando se cumple H_0 corresponde a: $s = (R-1) + (C-1)$.

Sea \hat{E}_{ij} , para $i = 1, \dots, R$ y $j = 1, \dots, C$, el estimador de máxima verosimilitud del número esperado de observaciones que son clasificadas en la celda $[i, j]$ de la tabla de contingencia cuando se cumple la hipótesis nula. Para este problema, el estadístico Q definido en (2.13), corresponde a:

$$Q = \sum_{i=1}^R \sum_{j=1}^C \frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}. \quad (2.18)$$

Como la tabla de contingencia está conformada por RC celdas, y como $s = R + C - 2$ parámetros deben ser estimados cuando la hipótesis H_0 es verdadera, se tiene que bajo esta hipótesis, y cuando $n \rightarrow \infty$, (2.18) converge en distribución a una χ^2 con $(R-1)(C-1)$ grados de libertad.

Por otra parte, el número esperado de observaciones clasificadas en la celda (i, j) es np_{ij} . Cuando se cumple H_0 , se tiene que $p_{ij} = p_{i\bullet} p_{\bullet j}$. Por lo tanto, si $\hat{p}_{i\bullet}$ y $\hat{p}_{\bullet j}$ denotan los estimadores de máxima verosimilitud para $p_{i\bullet}$ y $p_{\bullet j}$ respectivamente, se cumple que $\hat{E}_{ij} = n\hat{p}_{i\bullet}\hat{p}_{\bullet j}$. Como $p_{i\bullet}$ corresponde a la probabilidad de que una observación sea clasificada en la i -ésima fila de la tabla, $\hat{p}_{i\bullet}$ corresponde a la proporción del total de observaciones que son clasificadas en la i -ésima fila; es decir, $\hat{p}_{i\bullet} = \frac{N_{i\bullet}}{n}$. Similarmente, $\hat{p}_{\bullet j} = \frac{N_{\bullet j}}{n}$, y se tiene que:

$$\hat{E}_{ij} = n \left(\frac{N_{i\bullet}}{n} \right) \left(\frac{N_{\bullet j}}{n} \right) = \frac{N_{i\bullet} N_{\bullet j}}{n}. \quad (2.19)$$

Sustituyendo el valor de \hat{E}_{ij} en (2.18), es posible calcular el valor del estadístico Q mediante los valores observados N_{ij} . De forma análoga a lo explicado anteriormente, la hipótesis H_0 es rechazada si $Q \geq c$, con c una constante apropiada.

Test de Homogeneidad

De forma análoga al test de independencia presentado, las tablas de contingencia pueden ser utilizadas para desarrollar un test de homogeneidad [27], [30]. En específico, interesa la problemática en donde se toman observaciones de diferentes poblaciones y se obtiene una variable aleatoria (unidimensional) para cada una de estas observaciones. Se desea estudiar la hipótesis de que la distribución de dicha variable aleatoria es la misma para cada una de las poblaciones estudiadas.

Para esto, se asume que las muestras aleatorias son tomadas de R poblaciones diferentes, representadas por las R filas de la tabla. Mientras que cada una de las observaciones obtenidas, en cada una de las poblaciones, puede ser clasificada como uno de C tipos posibles, representados por las C columnas de la tabla. De esta forma los datos obtenidos en las R poblaciones pueden ser representados por una tabla de contingencia de dimensión $R \times C$, similar a la necesaria para realizar un test de independencia anteriormente descrito. Se denota por p_{ij} , con $(i = 1, \dots, R)$ y $(j=1, \dots, C)$, la probabilidad de que una muestra aleatoria tomada de la i -ésima población sea clasificada como perteneciente al j -ésimo tipo. Con esto:

$$\sum_{j=1}^C p_{ij} = 1 \quad \text{para } i = 1, \dots, R.$$

Las hipótesis a ser testeadas corresponden a:

$$H_0 : p_{1j} = p_{2j} = \dots = p_{Rj} \quad \text{para } j = 1, \dots, C. \quad (2.20)$$

$$H_1 : \text{no se cumple } H_0. \quad (2.21)$$

La hipótesis nula presentada en (2.20) establece que la distribución de cada uno de los C tipos es la misma en las R poblaciones estudiadas. Cuando esta hipótesis es verdadera, la combinación de las R poblaciones tiene como resultado la generación de una población homogénea con respecto a la variable aleatoria referente a la clasificación de cada observación.

Se denota por $n_{i\bullet}$, para $(i = 1, \dots, R)$, el número de observaciones en la muestra aleatoria pertenecientes a la i -ésima población. Para $(j = 1, \dots, C)$, se denota por N_{ij} el número de observaciones en dicha población que son clasificadas como pertenecientes al j -ésimo tipo. De esta forma:

$$\sum_{j=1}^C N_{ij} = n_{i\bullet} \quad \text{para } i = 1, \dots, R.$$

Si n denota el número total de observaciones en las R poblaciones, y $N_{\bullet j}$ denota el número total de observaciones del tipo j en las R poblaciones, entonces se cumplen las relaciones (2.14) y (2.15).

Para desarrollar el test para las hipótesis (2.20) y (2.21), se asume que los valores de p_{ij} son conocidos y se considera el siguiente estadístico calculado con los valores de las observaciones de la i -ésima población:

$$\sum_{j=1}^C \frac{(N_{ij} - N_{i\bullet} p_{ij})^2}{N_{i\bullet} p_{ij}}.$$

Este estadístico corresponde al utilizado por el test de bondad de ajuste presentado en (2.10), para una muestra aleatoria conformada por $N_{i\bullet}$ observaciones tomadas de la i -ésima población. Por lo tanto, cuando el tamaño de la muestra $N_{i\bullet}$ es grande, la distribución del estadístico presentado se aproxima a la de una χ^2 con $C - 1$ grados de libertad.

Si se suma este estadístico para las R poblaciones, se obtiene el siguiente estadístico:

$$\sum_{i=1}^R \sum_{j=1}^C \frac{(N_{ij} - N_{i\bullet} p_{ij})^2}{N_{i\bullet} p_{ij}}. \quad (2.22)$$

Debido a que las observaciones tomadas en las R poblaciones son independientes, el estadístico (2.22) distribuye como la suma de R variables aleatorias independientes, en donde cada una de ellas distribuye asintóticamente como una distribución χ^2 con $C - 1$ grados de libertad. Por lo tanto, la distribución de (2.22) distribuye asintóticamente como una χ^2 con $R(C - 1)$ grados de libertad.

Sin embargo, los valores de las probabilidades P_{ij} son desconocidos, y sus valores deben ser estimados mediante observaciones tomadas en las R poblaciones. Cuando H_0 es verdadera, las observaciones tomadas de las R poblaciones, son tomadas de la misma distribución. Por lo tanto, el estimador de máxima verosimilitud de la probabilidad de que una de estas observaciones sea del j -ésimo tipo corresponde a la proporción del total de observaciones tomadas de las R poblaciones que son del tipo j . Es decir, el estimador de máxima verosimilitud para p_{ij} es el mismo para todos los valores de i ($i = 1, \dots, R$), y corresponde a $\hat{p}_{ij} = \frac{N_{\bullet j}}{n}$. Al reemplazar este estimador en (2.22), se obtiene:

$$Q = \sum_{i=1}^R \sum_{j=1}^C \frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}. \quad (2.23)$$

En donde:

$$\hat{E}_{ij} = \frac{N_{i\bullet} N_{\bullet j}}{n}. \quad (2.24)$$

Debido a que la distribución en las R poblaciones es la misma cuando la hipótesis nula es verdadera, y como $\sum_{j=1}^C p_{ij} = 1$ para dicha distribución en común, se deben estimar $C - 1$ parámetros en este problema. Por lo tanto, el estadístico Q , definido en (2.23), distribuye asintóticamente como una χ^2 con $R(C - 1) - (C - 1) = (R - 1)(C - 1)$ grados de libertad.

2.5. Desigualdades de Concentración

En teoría de probabilidades, las desigualdades de concentración [32] otorgan cotas probabilísticas de cómo una variable aleatoria se desvía de un cierto valor. La ley de los grandes número establece que la suma de variables aleatorias independientes, bajo supuestos débiles, es muy parecida a su esperanza con una alta probabilidad. Esta suma corresponde uno de los ejemplos más básicos de una desigualdad de concentración, conformada por una variable aleatoria concentrada en torno a su media.

La más básica de estas desigualdades corresponde a la desigualdad de Markov: para toda variable aleatoria no negativa \mathbf{X} , y $t > 0$,

$$\mathbb{P}\{\mathbf{X} \geq t\} \leq \frac{\mathbb{E}\{\mathbf{X}\}}{t}. \quad (2.25)$$

Una de las implicancias directas de (2.25) es que si $\phi(\cdot)$ es una función no negativa monotónicamente creciente, entonces para para cualquier variable aleatoria \mathbf{X} y escalar t , se tiene:

$$\mathbb{P}\{\mathbf{X} \geq t\} = \mathbb{P}\{\phi(\mathbf{X}) \geq \phi(t)\} \leq \frac{\phi(\mathbb{E}\{\mathbf{X}\})}{\phi(t)}.$$

Un caso de particular interés es cuando $\phi(x) = x^2$: si \mathbf{X} es una variable aleatoria cualquiera y $t > 0$; entonces,

$$\mathbb{P}\{|\mathbf{X} - \mathbb{E}\{\mathbf{X}\}|^2 \geq t^2\} \leq \frac{\mathbb{E}\{|\mathbf{X} - \mathbb{E}\{\mathbf{X}\}|^2\}}{t^2}. \quad (2.26)$$

Utilizando el hecho de que $\mathbb{P}\{|\mathbf{X} - \mathbb{E}\{\mathbf{X}\}| \geq t\} = \mathbb{P}\{|\mathbf{X} - \mathbb{E}\{\mathbf{X}\}|^2 \geq t^2\}$, junto con la definición de la varianza, se obtiene la desigualdad de Chebyshev:

$$\mathbb{P}\{|\mathbf{X} - \mathbb{E}\{\mathbf{X}\}| \geq t\} \leq \frac{Var\{\mathbf{X}\}}{t^2}. \quad (2.27)$$

Sea la variable aleatoria $S_n = \sum_{i=1}^n \mathbf{X}_i$, en donde $\mathbf{X}_1, \dots, \mathbf{X}_n$ son variables aleatorias reales e independientes. Mediante la desigualdad de Chebyshev, y haciendo uso de la propiedad de independencia, se obtiene:

$$\mathbb{P}\{|S_n - \mathbb{E}\{S_n\}| \geq t\} \leq \frac{Var\{S_n\}}{t^2} = \frac{\sum_{i=1}^n Var\{\mathbf{X}_i\}}{t^2}.$$

De forma análoga, cuando se considera la variable aleatoria $\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i$, en donde $\mathbf{X}_1, \dots, \mathbf{X}_n$ son variables aleatorias reales e independientes, se obtiene para $t > 0$ la siguiente desigualdad:

$$\mathbb{P}\left\{\left|\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i - \mathbb{E}\{\mathbf{X}_i\}\right| \geq t\right\} \leq \frac{\sigma^2}{nt^2}. \quad (2.28)$$

En donde,

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n \text{Var}\{\mathbf{X}_i\}.$$

Una de las desigualdades de concentración muy utilizadas corresponde a la desigualdad de Hoeffding [33], la cual para las variables aleatorias \bar{X} y S_n anteriormente definidas, establece que: sean $\mathbf{X}_1, \dots, \mathbf{X}_n$ variables aleatorias independientes y acotadas; es decir, $\mathbb{P}\{\mathbf{X}_i \in [a_i, b_i]\} = 1$ ($i = 1, \dots, n$). Entonces; para cualquier $t > 0$,

$$\mathbb{P}\{|S_n - \mathbb{E}\{S_n\}| \geq t\} \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \quad (2.29)$$

$$\mathbb{P}\{|\bar{X} - \mathbb{E}\{\bar{X}\}| \geq t\} \leq 2 \exp\left(-\frac{2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \quad (2.30)$$

2.6. Cadenas de Markov

Un sistema cuyo estado evoluciona en el tiempo de forma aleatoria puede ser descrito mediante un proceso estocástico. En el caso de que el estado del sistema se observe en instantes de tiempo discretos, que se asumen iguales a \mathbb{N} ; y cuyo valor corresponde a un conjunto finito y numerable, se considera que el sistema evoluciona mediante un proceso estocástico en tiempo discreto definido en un conjunto finito. Se denota por \mathbf{X}_n a dicho estado del sistema en el tiempo $n \in \mathbb{N}$, el cual corresponde a una variable aleatoria definida en algún conjunto finito $S = \{S_1, \dots, S_r\}$. De esta forma la evolución temporal del sistema puede ser representada por la secuencia de variables aleatorias $\{\mathbf{X}_n | n \in \mathbb{N}\}$.

Se denota por $\pi_i(n)$ la probabilidad *a priori* de que el estado del sistema en el tiempo n sea E_i ; es decir, $\mathbb{P}\{\mathbf{X}_n = S_i\} = \pi_i(n)$. De esta forma, el vector $\pi(n) = [\pi_1(n), \dots, \pi_r(n)]$ corresponde a la distribución de probabilidad asociada a \mathbf{X}_n .

Para describir probabilísticamente un proceso estocástico en tiempo discreto, se debe conocer la ley inicial del sistema, denotada por $\pi(0)$, que corresponde a la distribución de probabilidad de \mathbf{X}_0 ; junto con la ley condicional de \mathbf{X}_n , que de forma recursiva describe la distribución de probabilidad de \mathbf{X}_n condicional a los estados del sistema en los instantes pasados. Es decir,

$$\mathbb{P}\{\mathbf{X}_n = S_{i_n} | \mathbf{X}_0 = S_{i_0}, \mathbf{X}_1 = S_{i_1}, \dots, \mathbf{X}_{n-1} = S_{i_{n-1}}\} \quad \forall (i_0, \dots, i_n) \in \{1, \dots, r\}^{n+1}.$$

Uno de los procesos estocásticos ampliamente estudiados, debido a su generalidad para representar diversas situaciones y su simpleza para realizar un estudio detallado, corresponde a las cadenas de Markov [34], [35]. En particular, se describirán las cadenas de Markov en tiempo discreto y definidas en un conjunto finito. Para esto, se define la propiedad de Markov débil; la cual establece

que la distribución de probabilidad de \mathbf{X}_{n+1} , dado que se conoce los estados previos del sistema ($\mathbf{X}_0 = S_{i_0}, \mathbf{X}_1 = S_{i_1}, \dots, \mathbf{X}_n = S_{i_n}$), depende sólo del último estado S_{i_n} alcanzado por éste:

$$\mathbb{P}\{\mathbf{X}_{n+1} = S_{i_{n+1}} | \mathbf{X}_0 = S_{i_0}, \mathbf{X}_1 = S_{i_1}, \dots, \mathbf{X}_n = S_{i_n}\} = \mathbb{P}\{\mathbf{X}_{n+1} = S_{i_{n+1}} | \mathbf{X}_n = S_{i_n}\}. \quad (2.31)$$

Un proceso estocástico de tiempo discreto y definido en un conjunto finito que cumple con la propiedad de Markov es conocido como una cadena de Markov discreta y finita de primer orden. El término primer orden se debe a la cantidad de estados sucesivos ya alcanzados por el sistema que influyen en el estado en que estará el mismo en el siguiente instante de tiempo. Así, una cadena de primer orden corresponde a un proceso en donde el estado del sistema en el siguiente instante de tiempo sólo depende de su estado actual, siendo independiente de la trayectoria de estados previa. De forma análoga es posible definir cadenas de Markov de orden $n \in \mathbb{N}$, en donde $n = 0$ corresponde a un proceso completamente independiente de su trayectoria de estados.

Para el caso de cadenas de Markov, la descripción probabilística del proceso queda totalmente especificada por el vector de probabilidades inicial del sistema: $\pi(0) = [\pi_1(0), \dots, \pi_r(0)]$; y la ley de evolución del sistema, correspondiente a la distribución de probabilidad de \mathbf{X}_n , condicionada a \mathbf{X}_{n-1} . Es decir, conocer la matriz de probabilidades de transición:

$$P_n^{n-1}[i, j] = \mathbb{P}\{\mathbf{X}_n = S_j | \mathbf{X}_{n-1} = S_i\} \quad \forall n \in \mathbb{N}, \quad \forall (i, j) \in \{1, \dots, r\}^2. \quad (2.32)$$

La distribución de probabilidad *a priori* para el estado del sistema en cualquier instante de tiempo queda determinada por:

$$\pi(n) = \pi(0) \prod_{i=1}^n P_i^{i-1}. \quad (2.33)$$

Un caso particular corresponde cuando la probabilidad de transición desde el estado S_i en el instante n hacia el estado S_j en el instante $n + k$ es independiente de n . Esto se traduce en que las matrices de probabilidades de transición sean constantes; es decir, $P_n^{n-1} = P \quad \forall n$. En tal caso se dice que el proceso corresponde a una cadena de Markov homogénea, y la ecuación (2.33) se reduce a:

$$\pi(n) = \pi(0)(P)^n. \quad (2.34)$$

En donde $(P)^n$ corresponde a la potencia de orden n de la matriz P .

En muchos casos es importante estudiar el comportamiento a largo plazo del sistema, en concreto interesa el estudio del vector $\lim_{n \rightarrow +\infty} \pi(n)$. Para el caso de una cadena de Markov finita y homogénea con matriz de probabilidades de transición P , se dice que π es un vector de probabilidades estacionarias si independiente de la condición inicial $\pi(0)$ se cumple que:

$$\pi = \lim_{n \rightarrow +\infty} \pi(n).$$

Cada una de las componentes del vector de probabilidades estacionarias π_j puede ser interpretado como la fracción del tiempo, luego de que el sistema haya evolucionado durante un número considerable de períodos, de que el sistema se encuentra en cada uno de los estados S_j .

2.7. Inferencia Bayesiana y Métodos de Monte Carlo

Un problema común en diversas disciplinas científicas; tales como estadística aplicada, procesamiento estadístico de señales, análisis de series de tiempo y econometría, corresponde a la estimación de variables no observables, utilizando para ello un conjunto de variables que sí son posibles de ser medidas. Para la problemática específica desarrollada en este trabajo; las variables no observables corresponden al SOC, junto con algunos parámetros internos del ESD. Mientras que las variables observadas corresponden al voltaje en bornes y la corriente presentes en el ESD durante su proceso de descarga.

Por otro lado, si bien es muy útil contar con una adecuada estimación del SOC, para muchas aplicaciones reales el conocimiento de dicha variable no es suficiente, ya que además se necesita tener una estimación de su comportamiento en el largo plazo. Por ejemplo; en el caso de un vehículo eléctrico (EV, por sus siglas en inglés) además de necesitarse el conocimiento de la cantidad de energía que hay disponible en su batería, es crucial poder saber la autonomía que se dispone para poder realizar un determinado recorrido. Esto corresponde a la etapa de pronóstico del SOC, la cual requiere haber realizado una adecuada etapa de estimación para dicha variable.

2.7.1. Inferencia Bayesiana

La evolución de un sistema dinámico puede ser descrita mediante su formulación en un modelo de espacio-estado. Este modelo corresponde a las ecuaciones (2.35) y (2.36); las cuales describen tanto la evolución del vector de estados del sistema $\{x_k, k \in \mathbb{N}\}$, como la evolución del vector de observaciones del mismo $\{z_k, k \in \mathbb{N}\}$.

$$x_k = f_k(x_{k-1}, v_{k-1}) \quad (2.35)$$

$$z_k = h_k(x_k, n_k); \quad (2.36)$$

en donde, $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_x}$ y $h_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_n} \rightarrow \mathbb{R}^{n_z}$ son funciones no lineales; $\{v_{k-1}, k \in \mathbb{N}\}$ y $\{n_k, k \in \mathbb{N}\}$ corresponden a secuencias i.i.d. (independiente e idénticamente distribuido) de ruido de proceso y medición, respectivamente; n_x y n_v son las dimensiones de los vectores de estado y ruido de proceso, respectivamente; n_z y n_n son las dimensiones de los vectores de observación y el ruido de observación, respectivamente.

Sea $z_{1:k} = \{z_i, i = 1, \dots, k\}$ el conjunto de todas las observaciones disponibles hasta el instante k . De igual forma, se define $x_{0:k} = \{x_i, i = 0, \dots, k\}$ como la trayectoria del vector de estados hasta el instante k . El problema de filtrado Bayesiano, o filtrado óptimo, consiste en obtener de forma recursiva algún grado de conocimiento sobre el estado x_k , utilizando para ello las observaciones $z_{1:k}$ disponibles en el instante k . Es decir, se requiere conocer la función de densidad de probabilidad (pdf, por sus siglas en inglés) de $p(x_k|z_{1:k})$, conocida también como la distribución *a posteriori* del vector de estados x_k . Una vez conocida esta distribución, es posible obtener estimaciones de ella, tales como su moda o media. Se asume que la condición inicial del vector de estados, $p(x_0|z_0) \equiv p(x_0)$, es conocida. Con esto, la pdf $p(x_k|z_{1:k})$ puede ser obtenida, de forma recursiva,

mediante dos etapas: predicción y actualización.

Suponiendo que la pdf requerida en el instante $k - 1$, $p(x_{k-1}|z_{1:k-1})$, se encuentra disponible. La etapa de predicción consiste en, utilizando la ecuación de transición de estados del modelo (2.35), obtener la pdf *a priori* del vector de estados en el instante k mediante la ecuación de Chapman-Kolmogorov:

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1}p(x_{k-1})|z_{1:k-1})dx_{k-1}. \quad (2.37)$$

La etapa de actualización consiste en, haciendo uso de una nueva observación que se obtiene en el instante k , modificar la distribución *a priori*, con el objetivo de obtener la densidad *a posteriori* en el instante k , mediante el teorema de Bayes [36]:

$$p(x_k|z_{1:k}) = \frac{p(z_{1:k}|x_k)p(x_k)}{p(z_{1:k})} \quad (2.38)$$

$$= \frac{p(z_k, z_{1:k-1}|x_k)p(x_k)}{p(z_k, z_{1:k-1})} \quad (2.39)$$

$$= \frac{p(z_k|z_{1:k-1}, x_k)p(z_{1:k-1}|x_k)p(x_k)}{p(z_k|z_{1:k-1})p(z_{1:k-1})} \quad (2.40)$$

$$= \frac{p(z_k|z_{1:k-1}, x_k)p(x_k|z_{1:k-1})p(z_{1:k-1})p(x_k)}{p(z_k|z_{1:k-1})p(z_{1:k-1})p(x_k)} \quad (2.41)$$

$$= \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}. \quad (2.42)$$

Así, se obtiene la siguiente relación recursiva:

$$\underbrace{p(x_k|z_{1:k})}_{\text{Posteriori}} = \frac{\overbrace{p(z_k|x_k)}^{\text{Verosimilitud}} \overbrace{p(x_k|z_{1:k-1})}^{\text{Priori}}}{\underbrace{p(z_k|z_{1:k-1})}_{\text{Evidencia}}}. \quad (2.43)$$

La constante de normalización en la ecuación (2.43),

$$p(z_k|z_{1:k-1}) = \int p(x_k|z_k)p(x_k|z_{1:k-1})dx_k, \quad (2.44)$$

depende de la función de verosimilitud $p(z_k|x_k)$, definida por la ecuación de observación del modelo (2.36), y la distribución *a priori* del vector de estados en el instante k (2.37).

Las relaciones recursivas (2.37) y (2.43) corresponden a la base de la solución Bayesiana óptima. Esta solución resuelve el problema de calcular la distribución *a posteriori* exacta en cada instante de tiempo, de forma recursiva. Sin embargo, esta propagación recursiva de la densidad *a posteriori* es sólo una solución conceptual, la cual no puede ser determinada en el caso general. Sólo en un selectivo conjunto de casos estas soluciones existen; incluyendo el Filtro de Kalman, para el caso de un modelo de espacio-estado lineal y Gaussiano [37], y Filtros Basados en Mallas [38].

2.7.2. Métodos Secuenciales de Monte Carlos: Filtro de Partículas

Los métodos secuenciales de Monte Carlo o Filtro de Partículas (PF, por sus siglas en inglés), corresponden a una técnica para la implementación del filtrado Bayesiano mediante simulaciones de Monte Carlo. La idea es representar la pdf *a posteriori* requerida por medio de un conjunto de muestras aleatorias, con pesos asociados, también llamadas partículas; y calcular estimadores mediante ellas. A medida que el número de muestras crece, esta caracterización se transforma en una descripción equivalente de la pdf *a posteriori*, y el Filtro de Partículas se aproxima a la solución Bayesiana óptima [38].

Sea $\{x_{0:k}^i, w_k^i\}_{i=1}^{N_s}$ un conjunto de partículas que caracteriza la pdf *a posteriori* $p(x_{0:k}|z_{1:k})$; en donde $\{x_{0:k}^i, i = 1, \dots, N_s\}$ es un conjunto de puntos de soporte, cuyos pesos asociados son $\{w_k^i, i = 1, \dots, N_s\}$, y $x_{0:k} = \{x_0, \dots, x_k\}$ corresponde a la trayectoria de los estados desde el instante 0 hasta el k . Los pesos son normalizados, tal que $\sum_{i=1}^{N_s} w_k^i = 1$. Con esto, la densidad *a posteriori* en el instante k puede ser aproximada como:

$$p(x_{0:k}|z_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_{0:k} - x_{0:k}^i). \quad (2.45)$$

Muestreo Secuencial de Importancia

La aproximación (2.45) corresponde a una distribución empírica que pretende aproximar la verdadera densidad *a posteriori*, $p(x_{0:k}|z_{1:k})$, y cuyos pesos son elegidos mediante el principio de muestreo de importancia [39]. Este principio de basa en lo siguiente. Suponiendo que $p(x) \propto \pi(x)$ es una densidad de probabilidad; de la cual es difícil tomar muestras, pero es posible de ser evaluada. Sean $x^i \sim q(x)$, $i = 1, \dots, N_s$ muestras que son fácilmente generadas desde una distribución $q(\cdot)$, llamada densidad de importancia. Con esto, una aproximación ponderada de la densidad $p(\cdot)$ está dada por:

$$p(x) \approx \sum_{i=1}^{N_s} w^i \delta(x - x^i); \quad (2.46)$$

en donde,

$$w^i \propto \frac{\pi(x^i)}{q(x^i)} \quad (2.47)$$

corresponde al peso normalizado de la i -ésima partícula.

Por lo tanto, si las muestras $x_{0:k}^i$ son tomadas de una distribución de importancia $q(x_{0:k}|z_{1:k})$, los pesos en la ecuación (2.45) se definen mediante (2.47) como:

$$w_k^i \propto \frac{p(x_{0:k}|z_{1:k})}{q(x_{0:k}|z_{1:k})}. \quad (2.48)$$

En cada iteración, es posible disponer de muestras que constituyan una aproximación de $p(x_{0:k-1}|z_{1:k-1})$, y se desea aproximar $p(x_{0:k}|z_{1:k})$ con un nuevo conjunto de muestras y pesos asociados. Si la densidad de importancia se escoge tal que:

$$q(x_{0:k}|z_{1:k}) = q(x_k|x_{0:k-1}, z_{1:k})q(x_{0:k-1}|z_{1:k-1}), \quad (2.49)$$

se pueden obtener muestras $x_{0:k}^i \sim q(x_{0:k}|z_{1:k})$, aumentando el vector de muestras existentes $x_{0:k-1}^i \sim q(x_{0:k-1}|z_{1:k-1})$ con un nuevo estado $x_k^i \sim q(x_k|x_{0:k-1}, z_{1:k})$. Para actualizar los pesos, se utilizan las siguiente relaciones:

$$p(x_{0:k}|z_{1:k}) = \frac{p(z_k|x_{0:k}, z_{1:k-1})p(x_{0:k}|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (2.50)$$

$$= \frac{p(z_k|x_{0:k}, z_{1:k-1})p(x_k|x_{0:k-1}, z_{1:k-1})p(x_{0:k-1}|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (2.51)$$

$$= \frac{p(z_k|x_k)p(x_k|x_{k-1})}{p(z_k|z_{1:k-1})}p(x_{0:k-1}|z_{1:k-1}) \quad (2.52)$$

$$\propto p(z_k|x_k)p(x_k|x_{k-1})p(x_{0:k-1}|z_{1:k-1}). \quad (2.53)$$

Sustituyendo (2.49) y (2.53) en (2.48), se obtiene la siguiente ecuación para la actualización de los pesos:

$$w_k^i \propto \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)p(x_{0:k-1}^i|z_{1:k-1})}{q(x_k^i|x_{0:k-1}^i, z_{1:k})q(x_{0:k-1}^i|z_{1:k})} \quad (2.54)$$

$$= w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_{1:k})}. \quad (2.55)$$

En el caso de que $q(x_k|x_{0:k-1}, z_{1:k}) = q(x_k|x_{k-1}, z_k)$, la densidad de importancia sólo depende de x_{k-1} y z_k . Esto es muy útil en el caso que se sólo se desee estimar la densidad $p(x_k|z_{1:k})$ en cada instante de tiempo. En tales casos, sólo es necesario almacenar x_k^i , por lo que la trayectoria $x_{0:k-1}^i$ y la historia $z_{1:k-1}$ es descartada. La ecuación (2.55) queda como:

$$w_k^i \propto w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)}. \quad (2.56)$$

Mientras que la densidad *a posteriori* $p(x_k|z_{1:k})$ puede ser aproximada como:

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_k - x_k^i); \quad (2.57)$$

en donde, los pesos son definidos mediante (2.56). A medida que $N_s \rightarrow \infty$, la aproximación (2.57) se aproxima a la verdadera densidad *a posteriori* $p(x_k|z_{1:k})$ [38].

De esta forma, el método de muestreo secuencial de importancia (SIS, por sus siglas en inglés), consiste en propagar recursivamente el conjunto de muestras, junto con sus pesos asociados, cada vez que una nueva observación es obtenida secuencialmente.

Remuestreo

Un problema habitual del algoritmo de muestreo secuencial de importancia presentado, es el fenómeno de degenerancia. Este fenómeno corresponde al hecho de que, luego de unas pocas iteraciones; todas las partículas, excepto una, tienen asociado un peso despreciable. La varianza de estos pesos sólo puede crecer en el tiempo [37], por lo que la presencia de degenerancia es inevitable. Esto implica que se deben destinar importantes recursos computacionales para la actualización de partículas cuya contribución a la aproximación (2.57) es insignificante. Una medida para este efecto corresponde a el tamaño efectivo de la muestra N_{eff} [39], definido como:

$$N_{eff} = \frac{N_s}{1 + Var\{w_k^{*i}\}}; \quad (2.58)$$

en donde, $w_k^{*i} = \frac{p(x_k^i | z_{1:k})}{q(x_k^i | x_{k-1}^i, z_k)}$ corresponde al peso verdadero. N_{eff} no es posible ser evaluado exactamente, por lo que se utiliza el siguiente estimador:

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s} (w_k^i)}, \quad (2.59)$$

en donde, w_k^i corresponden a los pesos normalizados obtenidos mediante (2.55). Siempre se cumple la condición $N_{eff} \leq N_s$, y un valor pequeño de N_{eff} indica una gran degenerancia. Este fenómeno corresponde a un efecto indeseable para la implementación del Filtro de Partículas, por lo que debe ser corregido. Dentro de los enfoques más simples está el de utilizar un valor muy grande para N_s ; sin embargo, esto es impracticable en casos de estudio reales.

Uno de los métodos utilizados para reducir el fenómeno de degenerancia corresponde a remuestreo. Este método se aplica cada vez que un determinado nivel de degenerancia se hace presente; es decir, cuando N_{eff} cae bajo cierto umbral N_T . La idea del remuestreo es eliminar partículas cuyo peso asociado sea pequeño, y concentrarse en aquellas cuyo peso sea grande. En la etapa de remuestreo se generan un nuevo conjunto de muestras $\{x_k^{i*}\}$, mediante N_s remuestreos (con reemplazo) de la aproximación (2.57); tal que:

$$\mathbb{P}(x_k^{i*} = x_k^j) = w_k^j. \quad (2.60)$$

De esta forma, el conjunto $\{x_k^{i*}\}$ corresponde a una muestra i.i.d. de la distribución de probabilidad (2.60); por lo tanto, los pesos son actualizados a $w_k^i = \frac{1}{N_s}$.

A pesar de que el método de remuestreo reduce el efecto de degenerancia, también introduce algunos efectos negativos. Por una parte, limita la posibilidad de paralelizar, ya que todas las partículas deben estar combinadas para su ejecución. Por otro parte, las partículas que poseen mayores pesos w_k^i son estadísticamente elegidas muchas veces. Esto lleva a una pérdida de diversidad entre las partículas, ya que su población está conformada por la repetición de muchas de ellas.

El siguiente pseudocódigo describe el algoritmo del Filtro de Partículas:

Algorithm 1 Generic Particle Filter

Require: $[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_s}, z_k]$
Ensure: $[\{x_k^i, w_k^i\}_{i=1}^{N_s}]$
for $i = 1 : N_s$ **do**
 Draw $x_k^i \sim q(x_k | x_{k-1}^i, z_k)$.
 Assign the particle a weight, w_k^i , according to (2.56).
end for
Calculate total weight: $t = \sum_{i=1}^{N_s} w_k^i$.
for $i = 1 : N_s$ **do**
 Normalize: $w_k^i = \frac{w_k^i}{t}$.
end for
Calculate \hat{N}_{eff} usign (2.59).
if $\hat{N}_{eff} < N_T$ **then**
 Resample: $[\{x_k^i, w_k^i\}_{i=1}^{N_s}] = RESAMPLE[\{x_k^i, w_k^i\}_{i=1}^{N_s}]$
end if

2.7.3. Esquema de Pronóstico Basado en Filtro de Partículas para Sistemas Dinámicos No-Lineales

Un esquema de pronóstico de fallas puede ser entendido como la generación de predicciones a largo plazo que describen el comportamiento de algún indicador de falla. Esto se realiza con el propósito de estimar la vida útil remanente de un componente que presenta algún modo de falla, tomando como condiciones iniciales las entregadas por la etapa de estimación.

Para la realización de la etapa de pronóstico, se requiere la presencia de al menos una característica que indique la dimensión o severidad del evento crítico. En caso de que esté disponible más de una característica, siempre es posible combinarlas con el propósito de obtener una sola señal. Luego, mediante la utilización de un modelo de espacio-estado, es posible estimar el comportamiento a lo largo del horizonte de predicción de dicha característica.

Para ilustrar lo anterior se considera la predicción de la pdf condicional de los estados, $\hat{p}(x_{t+k}^i | x_{t+k-1}^i)$, la cual describe la evolución de los estados en el instante futuro $t+k$, ($k = 1, \dots, p$) cuando la partícula x_{t+k}^i es utilizada como condición inicial. Asumiendo que los pesos actuales $\{w_t^i\}_{i=1}^{N_s}$ son una buena representación de la pdf de los estados en el instante t , entonces es posible aproximar la pdf de los estados en el instante $t+k$ utilizando la ley de probabilidades totales y el conjunto de pesos en el instante $t+k-1$, como lo muestra la siguiente ecuación:

$$\hat{p}(x_{t+k} | \hat{x}_{1:t+k-1}) \approx \sum_{i=1}^{N_s} w_{t+k-1}^i \cdot \hat{p}(x_{t+k}^i | \hat{x}_{t+k-1}^i), \quad \forall k \in \{1, \dots, p\}. \quad (2.61)$$

Para evaluar la aproximación (2.61), el peso de cada partícula debe ser modificado en cada instante del horizonte de predicción, con el objetivo de incorporar el hecho de que la no linealidad del proceso y su ruido asociado pueden cambiar la forma de la pdf de los estados a medida que pasa

el tiempo. Esto hace que sea necesario realizar una actualización de dichos pesos, la cual no puede depender de la adquisición de observaciones en el horizonte de predicción, contrario a lo hecho para la etapa de estimación. Por otro lado, antes de realizar la predicción en el siguiente instante de tiempo, un nuevo conjunto de partículas debe ser generado dentro del dominio de (2.61). Para resolver estas dificultades se describen dos enfoques, las cuales han sido desarrollos y validados en [7].

Generación de Predicciones a Largo Plazo: Primer Enfoque [7]

Este enfoque predice la evolución temporal de cada partícula mediante el cálculo sucesivo de la esperanza de la ecuación de actualización del vector de estados (2.35) para cada instante de tiempo futuro. Para el cálculo de dicha esperanza se considera el valor del estado asociado a esa partícula como condición inicial, como lo muestra la ecuación (2.62).

$$\hat{x}_{t+p}^i = \mathbb{E}\{f_{t+p}(\hat{x}_{t+p-1}^i, v_{t+p})\}. \quad (2.62)$$

Este primer enfoque para predicciones de largo plazo es simple desde el punto de vista de costo computacional. Sin embargo, hace el supuesto de que el error que se genera al considerar los pesos de las partículas invariante para instantes de tiempo futuros es despreciable con respecto a otras fuentes de incertidumbre (imprecisiones del modelo, supuestos sobre el ruido de proceso u observación, etc).

Bajo el supuesto de que las condiciones mencionadas se cumplen, (2.62) es considerada suficiente para extender las trayectorias del vector de estados $\hat{x}_{0:t+k}^i$, mientras sus pesos asociados son propagados en el tiempo sin realizar cambios. Mediante la realización de simulaciones se ha podido observar que este método ofrece resultados satisfactorios sobre el comportamiento del sistema en aplicaciones reales.

Generación de Predicciones a Largo Plazo: Segundo Enfoque [7]

El segundo enfoque para predicciones en el largo plazo, el cual se encuentra basado en el algoritmo del Filtro de Partículas Regularizado (RPF, por sus siglas en inglés), propone una solución al problema de representación de la incertidumbre en los instantes futuros de tiempo. Este enfoque es particularmente útil cuando el horizonte de predicción es grande. En lugar de recalculer los pesos de las partículas en cada instante de predicción, se propone representar la incertidumbre en instantes futuros mediante un remuestreo de la pdf de los estados predichos (2.61). De esta forma, la distribución de los estados en los instantes futuros está dada sólo por la posición de las partículas, y no por sus pesos asociados.

Para esto, se considera la aproximación discreta para la pdf predicha de los estados (2.63), en donde $K(\cdot)$ es un kernel o función de distribución; que puede corresponder a la pdf del ruido de

proceso, un kernel Gaussiano o una versión rescalada del kernel de Epanechnikov (2.65).

$$\hat{p}(x_{t+k}|\hat{x}_{1:t+k-1}) \approx \sum_{i=1}^{N_s} w_{t+k-1}^i \cdot K_h(x_{t+k} - \mathbb{E}\{x_{t+k}^i|\hat{x}_{t+k-1}^i\}). \quad (2.63)$$

$$K_h = \frac{1}{h^{n_x}} K\left(\frac{x}{h}\right), \quad h_{opt} = A \cdot N^{-\frac{1}{n_x+4}}, \quad A = (8c_{n_x}^{-1} \cdot (n_x + 4) \cdot (2\sqrt{\pi})^{n_x})^{\frac{1}{n_x+4}}. \quad (2.64)$$

$$K(x) = \begin{cases} \frac{n_x+4}{2c_{n_x}}(1 - \|x\|) & \text{si } \|x\| < 1 \\ 0 & \text{si no.} \end{cases} \quad (2.65)$$

En donde, c_{n_x} corresponde al volumen de la esfera unitaria en \mathbb{R}^{n_x} . Para representar la incertidumbre existente en (2.63) se genera una nueva población de partículas para el instante $t+k$ ($k \in \{1, \dots, p\}$), cuyos pesos son uniformes, mediante un proceso de remuestreo inverso [7].

Para evitar pérdida de diversidad en la población de partículas, se desarrolla un paso adicional inspirado en RPF. Para esto, se asume que la la matriz de covarianza del vector de estados \hat{S}_{t+k} corresponde a la matriz de covarianza empírica de \hat{x}_{t+k} ; y que se dispone de un conjunto de partículas, de pesos uniformes, para \hat{x}_{t+k-1} . Con esto, la eficiencia en el uso de kernels de Epanechnikov para aproximar una pdf, en el sentido de error medio cuadrático, es maximizada.

El algoritmo de regularización aplicado a realizar predicciones a largo plazo se presenta a continuación.

- Aplicar el procedimiento de remuestreo de la transformada inversa modificado. For $i = 1, \dots, N_s$: $w_{t+k}^i = \frac{1}{N_s}$.
- Calcular \hat{S}_{t+k} , la matriz de covarianza empírica de $\{\mathbb{E}\{x_{t+k}^i|\hat{x}_{t+k-1}^i\}, w_{t+k}^i\}_{i=1}^{N_s}$.
- Calcular \hat{D}_{t+k} tal que $\hat{D}_{t+k}^T \hat{D}_{t+k} = \hat{S}_{t+k}$.
- For $i = 1, \dots, N_s$: obtener $\varepsilon^i \sim K$, de la distribución de Epanechnikov y asignar $\hat{x}_{t+k}^{i*} = \hat{x}_{t+k}^i + h_{t+k}^{opt} \hat{D}_{t+k} \varepsilon^i$, donde h_{t+k}^{opt} es calculado en (2.64).
- For $i = 1, \dots, N_s$: $\hat{x}_{t+k}^i = \hat{x}_{t+k}^{i*}$.

Cabe mencionar que el supuesto de ruidos de proceso no correlacionados sólo fue hecho con el objetivo de reducir el esfuerzo computacional al realizar el proceso de remuestreo. De hecho, no hay restricciones teóricas para la aplicación de esta metodología en el caso de tener ruidos de proceso correlacionados.

Estimación y Caracterización Estadística de la Vida Útil Remanente

Uno de los objetivos principales de un algoritmo de pronóstico corresponde a la estimación de la pdf de la vida útil remanente del sistema, la cual está relacionada con la probabilidad de existir

una falla en los instantes de tiempo futuros. Esta probabilidad se obtiene mediante las predicciones a largo plazo de los estados del sistema, junto con información empírica sobre las condiciones críticas. Estas condiciones son típicamente especificadas en forma de umbrales para los indicadores de falla, definiendo lo que se conoce como zona de peligro [7]. En la práctica, la zona de peligro queda determinada por una pdf empírica basada en datos históricos de la ocurrencia del evento crítico.

La definición de la ocurrencia de una falla en un sistema, y su correspondiente RUL, depende de cada problema en específico. Para el caso de pronóstico del SOC, la ocurrencia de este evento corresponde al instante de tiempo en donde ocurre el fin de la descarga (EOD, por sus siglas en inglés), el cuál está caracterizado por el instante en que el voltaje en bornes de la batería cae por debajo de un valor crítico.

Así, la pdf del EOD puede ser obtenida mediante la ley de probabilidades totales, como se muestra en la ecuación (2.66).

$$\mathbb{P}(EOD = eod) = \sum_{i=1}^{N_s} \mathbb{P}(Failure|X = \hat{x}_{eod}^i) \cdot w_{eod}^i. \quad (2.66)$$

Capítulo 3

Desarrollo Metodológico

En este capítulo se describe cómo es posible modelar el perfil de corriente de descarga de una batería mediante una cadena de Markov homogénea. Además, se presenta una metodología, la cual corresponde al núcleo de este trabajo, que permite encontrar todos los parámetros de dicha cadena mediante fundamentos teóricos. Para esto, se hace uso de las herramientas matemáticas presentadas en el Capítulo 2.

3.1. Inferencia Estadística sobre Cadenas de Markov

Dentro de las cadenas de Markov, las que presentan un comportamiento homogéneo corresponden a las más simples de ser estudiadas. Si además se dispone de un conjunto apropiado de observaciones equidistantes temporalmente para dicho proceso, es posible realizar inferencias estadísticas sobre él. Por ejemplo, es posible estimar las probabilidades de transición o realizar test de hipótesis sobre ellas.

En lo que sigue se explica la forma de realizar una inferencia estadística sobre cadenas de Markov homogéneas, y se describirán los detalles para una cadena de primer orden. La extensión para otros ordenes es generalizable a partir de lo aquí explicado. Sean $i \in \{1, \dots, m\}$ los estados de esta cadena, y $t \in \{0, \dots, T\}$ los períodos de observación del proceso. Se denota por p_{ij} ($(i, j) \in \{1, \dots, m\}^2$) a la probabilidad de estar en el estado j en el período t , dado que se estuvo en el estado i en el período $t - 1$ ($\forall t \in \{1, \dots, T\}$).

3.1.1. Estimación de las Probabilidades de Transición

Una observación para una realización de la cadena de Markov consiste en la secuencia de los estados presentados en $t = 0, \dots, T$; los cuales se denotan por $i(0), \dots, i(T)$ respectivamente. Se asume conocido el estado inicial $i(0)$, por lo que existe m^T posibles secuencias de estados. Estas secuencias

corresponden a eventos mutuamente excluyentes cuya función de verosimilitud corresponde a:

$$f_T(\mathbf{n}|P) = \prod_{i=1}^m \prod_{j=1}^m p_{ij}^{n_{ij}}. \quad (3.1)$$

En donde \mathbf{n} corresponde a una matriz cuya entrada $n[i, j]$ ($(i, j) \in \{1, \dots, m\}^2$) indica el número total de transiciones desde el estado i hacia el estado j en la observación de la realización del proceso, mientras que P denota su matriz de probabilidades de transición.

Las probabilidades de transición p_{ij} ($(i, j) \in \{1, \dots, m\}^2$) pueden ser estimadas mediante el estimador de máxima verosimilitud [40]. Para esto se debe maximizar (3.1) con respecto a p_{ij} . Sin embargo, para que el resultado de este procedimiento sea coherente, se debe agregar las restricciones $p_{ij} \geq 0 \quad \forall (i, j) \in \{1, \dots, m\}^2$ y $\sum_{j=1}^m p_{ij} = 1 \quad \forall i \in \{1, \dots, m\}$. Así, la obtención de dicho estimador se traduce a la resolución de un problema de optimización con restricciones, cuya solución viene dada por:

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^m n_{ij}} \quad \forall (i, j) \in \{1, \dots, m\}^2. \quad (3.2)$$

3.1.2. Test de Hipótesis de que la Cadena es de un Orden Determinado

En esta sección se describirá un procedimiento para testear la hipótesis nula de que la observación de la realización del proceso corresponde a una cadena de Markov homogénea de primer orden, en contra de su hipótesis alternativa.

Para esto, se define una cadena de Markov homogénea de segundo orden; la cual corresponde a una generalización de una cadena de Markov de primer orden, como un proceso cuyo estado en el siguiente instante de tiempo depende de su estado actual y anterior. Se denota por p_{ijk} ($\forall (i, j, k) \in \{1, \dots, m\}^3$) a la probabilidad de estar en el estado k en el período t , dado que se estuvo en los estados i y j en los períodos $t - 2$ y $t - 1$ ($\forall t \in \{2, \dots, T\}$), respectivamente. Cabe mencionar que una cadena homogénea de primer orden es un caso particular de una cadena de segundo orden, una en donde las probabilidades de transición p_{ijk} no dependen del primer índice i . Por otro lado, una cadena de segundo orden puede ser representada como una cadena de primer orden más compleja (para mayor detalle consultar [41]).

Con lo anteriormente explicado, es posible escribir las hipótesis estadísticas antes mencionadas como:

$$H_0 : p_{1jk} = p_{2jk} = \dots = p_{mjk} = p_{jk} \quad \forall (j, k) \in \{1, \dots, m\}^2. \quad (3.3)$$

$$H_1 : \text{no se cumple } H_0. \quad (3.4)$$

De forma análoga a lo hecho para el caso de una cadena de primer orden, el Estimador de Máxima Verosimilitud de las probabilidades de transición de una cadena de Markov homogénea

de segundo orden corresponde a:

$$\hat{p}_{ijk} = \frac{n_{ijk}}{\sum_{k=1}^m n_{ijk}} \quad \forall (i, j, k) \in \{1, \dots, m\}^3. \quad (3.5)$$

En donde n_{ijk} corresponde al número total de transiciones en la observación de la realización del proceso que se inician en el estado i , luego pasan por el estado j , y finalmente terminan en el estado k .

Para testear las hipótesis presentadas en (3.3) y (3.4), se utiliza una tabla de dimensión $m \times m$, la cual tiene la misma apariencia que una tabla de contingencia, para representar las estimaciones de \hat{p}_{ijk} para un valor de j dado y $\forall (i, k) \in \{1, \dots, m\}^2$. Bajo la hipótesis nula $p_{ijk} = p_{jk} \quad \forall i \in \{1, \dots, m\}$, por lo que el test χ^2 de homogeneidad es adecuado [40]. Para testear esta hipótesis, se calcula el estadístico:

$$\chi_j^2 = \sum_{i=1}^m \sum_{k=1}^m n_{ij} \frac{(\hat{p}_{ijk} - \hat{p}_{jk})^2}{\hat{p}_{jk}}. \quad (3.6)$$

Si la hipótesis nula es verdadera, (3.6) converge en distribución a una χ^2 con $(m-1)^2$ grados de libertad.

Con los pasos realizados anteriormente, las hipótesis (3.3) y (3.4) pueden ser testeadas separadamente para cada valor de j . Sin embargo, también es posible testear la hipótesis conjunta de que $p_{ijk} = p_{jk} \quad \forall (i, j, k) \in \{1, \dots, m\}^3$. Para realizar este test, se debe calcular el siguiente estadístico:

$$\chi^2 = \sum_{j=1}^m \chi_j^2 = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m n_{ijk} \frac{(\hat{p}_{ijk} - \hat{p}_{jk})^2}{\hat{p}_{jk}}, \quad (3.7)$$

el cual converge en distribución a una χ^2 con $m(m-1)^2$ grados de libertad.

3.2. Conjunto de Datos Utilizados

Para desarrollar la metodología presentada en este trabajo se ha utilizado dos conjuntos de datos. En específico cada uno de los conjuntos está compuesto por mediciones simultáneas de corriente, voltaje y tiempo, correspondientes a pruebas experimentales del proceso de descarga de una celda de Ion-Litio. El tiempo de muestreo de dichas variables fue de 1 [seg], y la descarga fue realizada hasta que el voltaje en bornes de la celda llegara a 0.5 [V], comenzando con ella totalmente cargada (estado de carga al máximo).

El primer conjunto de datos (denotado por *Set #1*) corresponde al proceso de descarga de una celda de Ion-Litio 18650 (3.7 [V] y 2.4 [Ah]) mediante un perfil de uso que emula a el exhibido

por la batería que energiza a un robot terrestre de cuatro ruedas, cuya corriente máxima y mínima fueron definidas en 2.809 [A] y 1.619 [A], respectivamente. Este conjunto está conformado por 2920 observaciones sucesivas. El segundo conjunto (denotado por *Set #2*) corresponde a la descarga de una celda de Ion-Litio 18650 (3.7 [V] y 3 [Ah]) conformada por 3638 observaciones sucesivas, en donde el perfil de uso corresponde a la realización de una cadena de Markov homogénea de dos estados, los cuales representan cambios aleatorios entre niveles altos y bajos de corriente. Las probabilidades de transición para esta cadena son $p_{11} = p_{21} = 0.55$ y $p_{12} = p_{22} = 0.45$, y los estados se definen como el valor de la corriente de descarga (estado #1: 1[A], estado #2: 3[A]).

La Figura 3.1 muestra los gráficos de los conjuntos de datos utilizados.

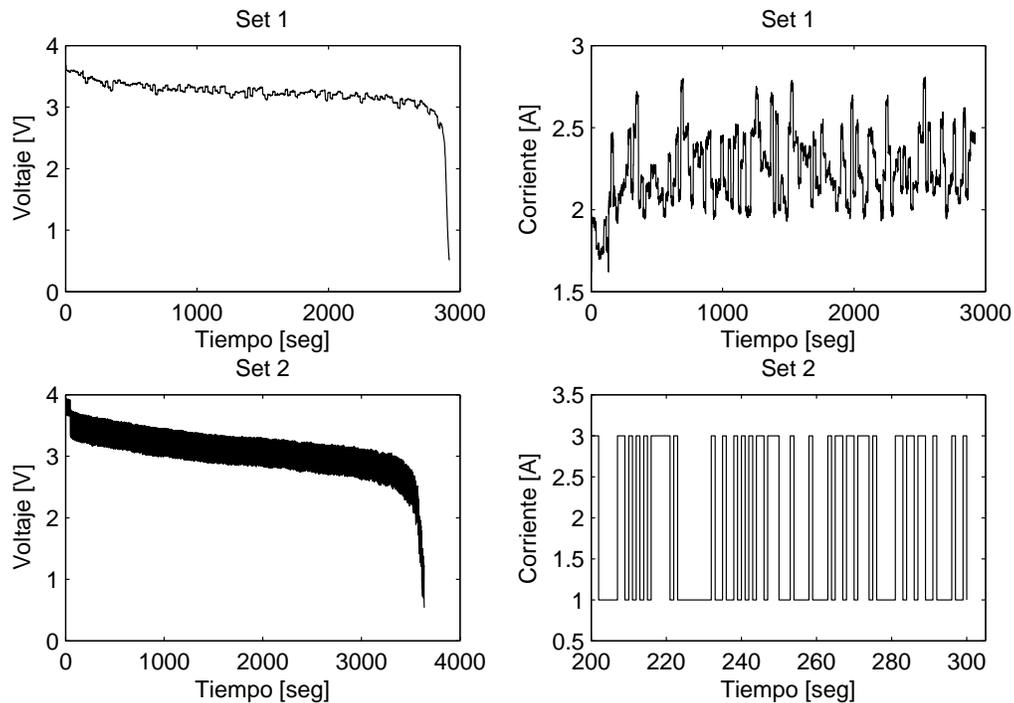


Figura 3.1: Conjunto de datos de voltaje y corriente correspondientes a pruebas experimentales de descarga de un ESD.

3.3. Elección de los Parámetros de la Cadena de Markov

Una cadena de Markov homogénea queda totalmente especificada por el valor que representan sus estados y la matriz de probabilidades de transición. A continuación se detallará la forma en que dichos parámetros pueden ser determinados a partir de la observación de una realización de dicha cadena.

3.3.1. Determinación del Valor de los Estados

Para determinar cuales son los estados de la cadena de Markov que representa el perfil de corriente utilizado durante la descarga de la batería, se hace uso del método de agrupamiento de datos mediante k-means. Para esto, se realiza un agrupamiento de los datos de dicho perfil correspondientes a la variable de la corriente, sin considerar la variable temporal, y definiendo el número de centroides requeridos *a priori*. De esta forma, cada una de estas mediciones de corriente tiene un centroide asociado, cuyo valor corresponde a su estado de la cadena de Markov.

Para ilustrar este procedimiento, la Figura 3.2 muestra la cadena de Markov, formada por tres estados, que representa al perfil de corriente del primer conjunto de datos (Set #1) en los primeros 1000 instantes de tiempo. En color negro se aprecia el perfil de corriente en función del tiempo, mientras que en color rojo se indica el valor del estado de la cadena de Markov al que pertenece cada muestra de dicho perfil de corriente.

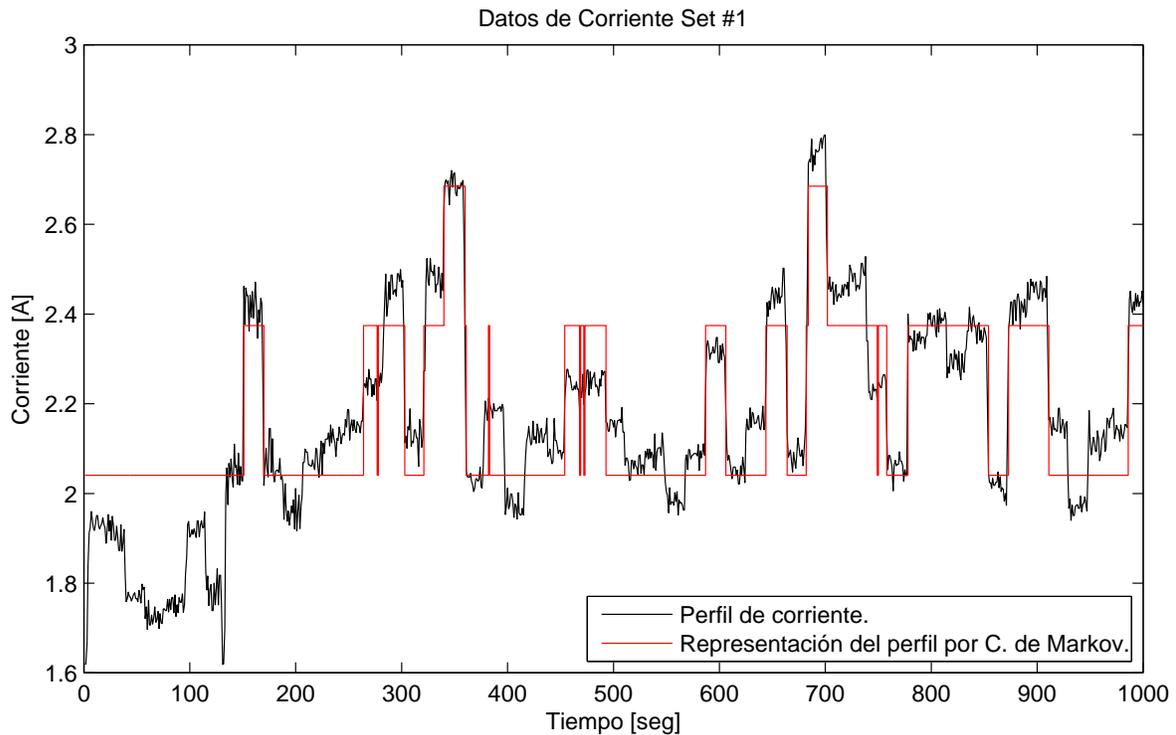


Figura 3.2: Datos de Corriente Set #1.

En la Figura 3.3, la cual corresponde a un acercamiento de la Figura 3.2, es posible apreciar dos tipos de situaciones no deseadas. Por un lado se observa que muestras contiguas del perfil de corriente que presentan un valor similar, tienen asociados un estado distinto en su representación mediante la cadena de Markov. Sin embargo, dichas muestras deben pertenecer a una misma agrupación, ya que describen un mismo punto de operación. Por otro lado, cuando el perfil de corriente cambia su valor abruptamente, se produce un cambio entre estados no colindantes de la cadena. Sin embargo, durante la transición pueden quedar unas pocas muestras cuyo estado

asociado es uno intermedio al de las muestras vecinas. Este último fenómeno se produce ya que al haber un gran cambio en el perfil de corriente, el ESD no es capaz de realizar dicho cambio en forma instantánea.

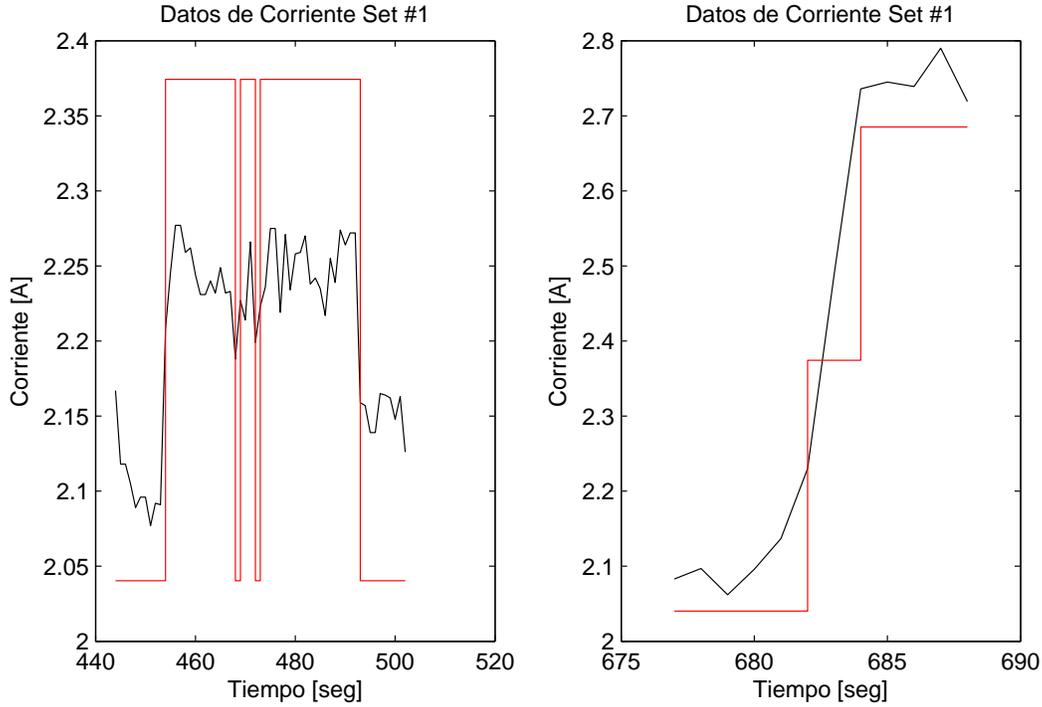


Figura 3.3: Acercamiento de Datos de Corriente Set #1, ilustrando dos situaciones irregulares.

Para solucionar las situaciones mencionadas, se debe realizar un pre-procesamiento de la señal correspondiente a la cadena de Markov. Para el primer caso presentado, el preprocesamiento consiste en permitir un cambio de estado sólo cuando el perfil de corriente presente un cambio de valor considerable. Mientras que para la segunda situación, el procedimiento consiste en no permitir que un pequeño conjunto de muestras contiguas, donde las muestras que lo anteceden y suceden pertenecen a estados no sucesivos, pertenezcan a un agrupamiento intermedio de dichos estados. La Figura 3.4 muestra el resultado de la aplicación de este preprocesamiento a los datos presentados en la Figura 3.2.

3.3.2. Determinación de las Probabilidades de Transición

Sea una cadena de Markov homogénea de orden $n \in \mathbb{N}$ conformada por m estados. Para determinar el valor de las $m^{(n+1)}$ probabilidades de transición asociadas a esta cadena, éstas deben ser estimadas mediante su estimador de máxima verosimilitud anteriormente presentado. Sean n índices, denotados por i_l ($l \in \{1, \dots, n\}$), cada uno de ellos definido en el conjunto $\{1, \dots, m\}$. La estimación de cada una de las probabilidades de transición viene dada por:

$$\hat{p}_{i_1 i_2 \dots i_n j} = \frac{n_{i_1 i_2 \dots i_n j}}{\sum_{j=1}^m n_{i_1 i_2 \dots i_n j}} \quad \forall (i_1, i_2, \dots, i_n, j) \in \{1, \dots, m\}^{(n+1)}. \quad (3.8)$$

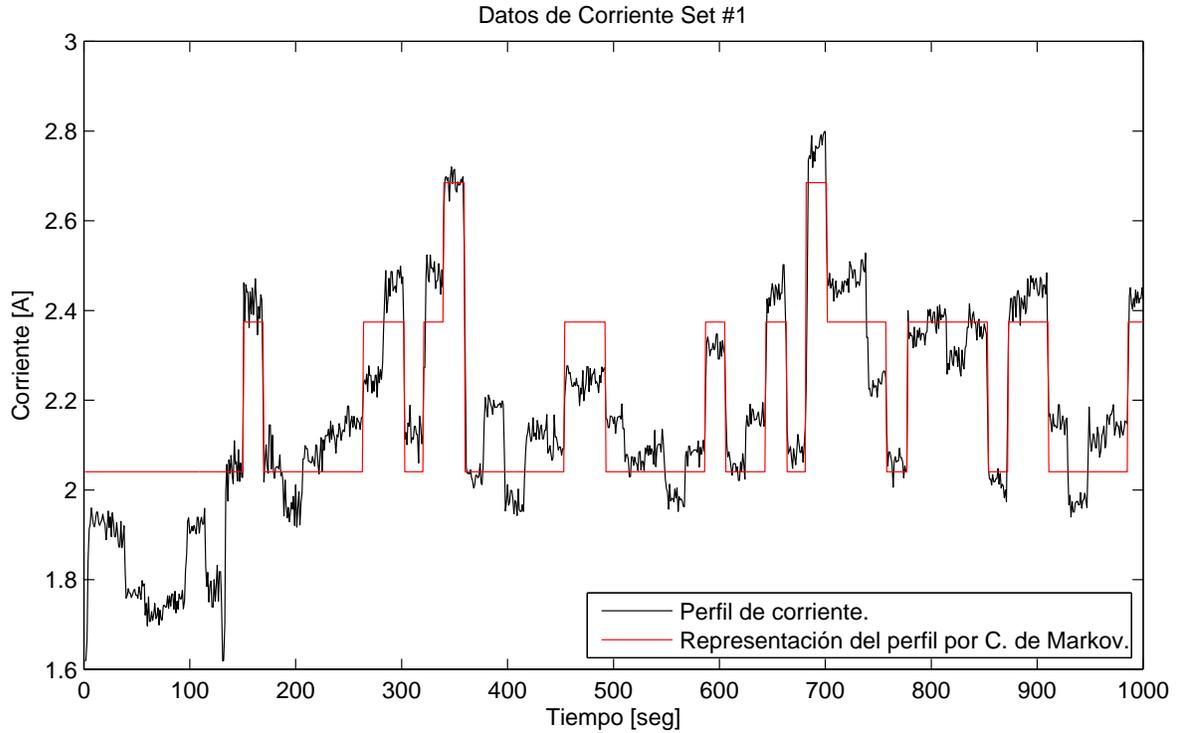


Figura 3.4: Preprocesamiento de la cadena de Markov correspondiente Datos de Corriente Set #1.

En donde $n_{i_1 i_2 \dots i_n j}$ corresponde al número total de transiciones, en la observación de la realización del proceso, que siguieron la trayectoria de estados consecutivos correspondiente a $i_1 i_2 \dots i_n j$.

3.3.3. Determinación de la Cantidad de Estados

Anteriormente se explicó cómo definir los estados de la cadena mediante el uso de k-means, asumiendo un número de centroides predeterminado. Mientras mayor sea el número de estados con los que se define la cadena, es de esperar que ésta represente de mejor forma el perfil de corriente asociado. Sin embargo, la elección del máximo número de estados posibles no es arbitraria. Esto se debe a que a medida que se define un mayor número de centroides, la cantidad de observaciones que pertenecen a él disminuye. Al tener un menor número de observaciones correspondientes a cada estado de la cadena, la estimación de las probabilidades de transición entre dichos estados es de menor calidad. Para cuantificar este efecto se utilizan las desigualdades de concentración ya presentadas. A continuación se describirán los detalles para una cadena de primer orden. La extensión para órdenes mayores es generalizable a partir del caso de primer orden.

Para esto se considera que cada una de las transiciones entre estados presentes en la observación de la realización de la cadena corresponde a una variable aleatoria multinomial, cuya distribución depende del estado en el que se produjo dicha transición y corresponde a las probabilidad de transición de la cadena.

En específico, si la realización de la cadena de Markov está conformada por m estados y hay n transiciones desde el estado $i \in \{1, \dots, m\}$, interesa estudiar la calidad de la estimación de p_{ij} ($j \in \{1, \dots, m\}$). Las transiciones observadas desde el estado i hacia el estado j pueden ser descritas mediante una secuencia de variables aleatorias \mathbf{X}_k (con $k \in \{1, \dots, n\}$), cada una de ellas correspondiente a un experimento de Bernoulli que toma el valor 1 (con probabilidad p_{ij}) si la transición efectivamente fue hacia el estado j , y 0 (con probabilidad $(1 - p_{ij})$) si no. Este último caso corresponde a una transición desde el estado i hacia un estado distinto de j . De esta forma, la ecuación (2.30) queda de la siguiente forma:

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{k=1}^n \mathbf{X}_k - \mathbb{E}\{X_1\} \right| \geq t \right\} \leq 2 \exp \left(-\frac{2n^2 t^2}{n} \right). \quad (3.9)$$

Debido a la elección de la secuencia de las variables aleatorias $\{\mathbf{X}_k, k \in \mathbb{N}\}$, la expresión $\frac{1}{n} \sum_{k=1}^n \mathbf{X}_k$ corresponde a $\frac{n_{ij}}{n}$; es decir, el estimador de máxima verosimilitud para la probabilidad de transición p_{ij} . Por otro lado, dicha secuencia está conformada por variables aleatorias independientes e idénticamente distribuidas entre sí, cada una de ellas con esperanza $\mathbb{E}\{\mathbf{X}_k\} = p_{ij}$ ($\forall k \in \{1, \dots, n\}$), por lo que la ecuación (3.9) se escribe como:

$$\mathbb{P} \{ |\hat{p}_{ij} - p_{ij}| \geq t \} \leq 2 \exp(-2nt^2). \quad (3.10)$$

También, la varianza de estas variables aleatorias corresponde a $Var\{\mathbf{X}_k\} = p_{ij}(1 - p_{ij})$ ($\forall k \in \{1, \dots, n\}$). Dado que $p_{ij} \in (0, 1)$, el máximo valor para la expresión $p_{ij}(1 - p_{ij})$ corresponde a $\frac{1}{4}$ y se obtiene cuando $p_{ij} = \frac{1}{2}$. Así, es posible obtener una cota superior para ésta varianza dada por $Var\{\mathbf{X}_k\} \leq \frac{1}{4}$ ($\forall k \in \{1, \dots, n\}$), la cual puede ser utilizada en la ecuación (2.28) para obtener la siguiente desigualdad:

$$\mathbb{P} \{ |\hat{p}_{ij} - p_{ij}| \geq t \} \leq \frac{1}{4nt^2}. \quad (3.11)$$

Las ecuaciones (3.10) y (3.11) especifican cotas superiores para la probabilidad de que el estimador de p_{ij} difiera de su valor real por un valor igual o mayor que t . Estas cotas, para un valor fijo de t , corresponden a funciones decrecientes de n , el número de transiciones observadas desde el estado i . De esta forma, a medida que se tiene una mayor cantidad de dichas observaciones, es posible obtener una mejor estimación para la probabilidad de transición. La Figura 3.5 ilustra el comportamiento de las cotas mencionadas, para $t = 0.1$, a medida que n aumenta.

De la Figura 3.5 se aprecia que hasta un cierto valor de n , el cual depende del valor elegido para t , las cotas superiores de la probabilidad es mayor a 1, por lo que no entregan ninguna información útil. Por otro lado, a partir de un determinado valor de n , que también depende de t , la cota entregada por la desigualdad (3.10) es menor a la entregada por (3.11). Estos valores de n se definen como:

$$n_1(t) = \sup_{n \in \mathbb{N}} \left\{ \frac{1}{4nt^2} \geq 1 \right\}. \quad (3.12)$$

$$n_2(t) = \sup_{n \in \mathbb{N}} \left\{ 2 \exp(-2nt^2) \geq \frac{1}{4nt^2} \right\}. \quad (3.13)$$

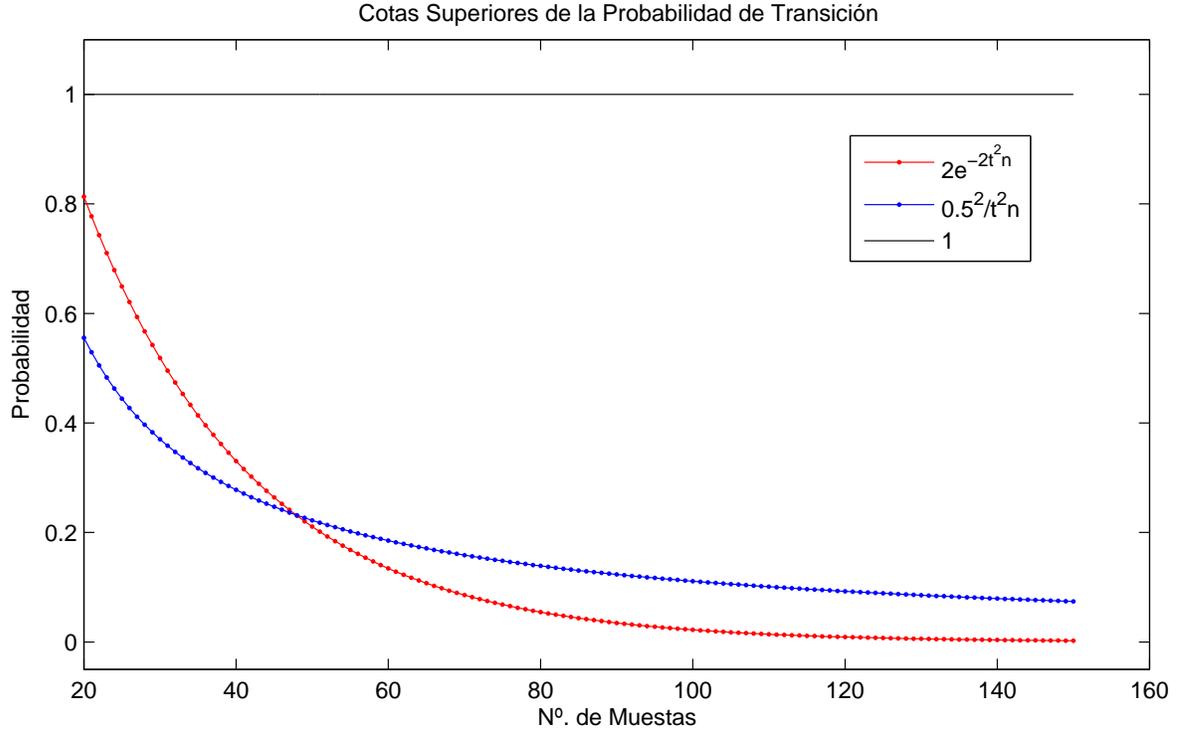


Figura 3.5: Comportamiento de las cotas superiores en función del número de observaciones, utilizando $t = 0.1$.

Con esto, se obtiene la siguiente cota superior para la probabilidad en función de n y para un valor dado de t :

$$c(n) = \begin{cases} 1 & \text{si } n \leq n_1(t) \\ 1/(4nt^2) & \text{si } n_1(t) < n \leq n_2(t) \\ 2 \exp(-2nt^2) & \text{si } n_2(t) < n \end{cases} \quad (3.14)$$

La cota $c(n)$ se utiliza para determinar el máximo número de estados que posee la cadena de Markov. Para esto se escoge un valor de $t \in (0, 1)$ junto con un escalar $p^* \in (0, 1)$, los cuales corresponden a parámetros de diseño. Estos valores están relacionados, mediante la desigualdad (3.15), con la máxima probabilidad que se está dispuesto a aceptar de que el estimador \hat{p}_{ij} difiera de su valor real p_{ij} en un valor mayor o igual a t .

$$\mathbb{P} \{ |\hat{p}_{ij} - p_{ij}| \geq t \} \leq p^*. \quad (3.15)$$

Para obtener las estimaciones requeridas para p_{ij} ($\forall j \in \{1, \dots, m\}$), se debe cumplir que $c(n) \leq p^*$. Esta desigualdad se debe cumplir para todos los estados $i \in \{1, \dots, m\}$, de modo que las m^2 probabilidades de transición sean estimadas adecuadamente. El hecho de que dicha desigualdad no se cumpla para algún estado i , se debe a que no hay suficientes transiciones observadas, que partan en dicho estado, para cumplir los requerimiento de la estimación. En estos casos, se debe reducir el número de estados con que se representa la cadena, para que de esta forma haya un mayor número de muestras que pertenecen a cada estado. Este procedimiento de reducción del número de estados se realiza en forma progresiva hasta que se logre obtener el mayor número de ellos con los que se cumpla el requerimiento de estimación, o hasta llegar a representar la cadena con un solo estado. En este último caso el valor del estado corresponde al promedio del perfil de corriente.

3.4. Elección del Orden de la Cadena de Markov

Una vez que se ha explicado cómo determinar la cantidad de estados de la cadena, junto con el valor de los mismos, y las probabilidades de transición entre ellos; hay que comprobar que dicha cadena sea de un orden determinado.

El primer paso que se debe realizar para llevar a cabo este propósito consiste en determinar cuál es el mayor orden de una cadena de Markov con que se puede representar el perfil de corriente, utilizando en cada uno de ellos el mayor número de estados posibles. La forma de obtener este mayor orden (denotado por O_{max}) consiste en, partiendo desde $O_{max} = 1$, verificar si es posible obtener una cadena de Markov de dicho orden con un número de estados mayor a 1. De ser esto posible, se aumenta progresivamente el valor de O_{max} en una unidad y se vuelve a verificar la condición presentada, hasta obtener un valor para el orden de la cadena (denotado por O_{fin}) en donde el máximo número de estados sea 1. Finalmente, $O_{max} = O_{fin} - 1$.

Una vez determinado el valor de O_{max} , se prueban la generalización de las Hipótesis (3.3) y (3.4), partiendo desde el orden 0; y considerando el mayor número de estados posibles. Cabe mencionar que para dicho orden, estas hipótesis corresponden a las utilizadas en el test de independencia anteriormente presentado. El mayor número de estados posibles para la cadena de orden 0, corresponde a el encontrado para la cadena de orden 1. De cumplirse H_0 , se ha logrado obtener el orden de la cadena de Markov requerida. En caso contrario, se aumenta progresivamente el orden de la cadena, volviendo a testear dichas hipótesis estadísticas, hasta llegar al orden O_{max} . Si no se cumple H_0 para este último orden, se reduce en una unidad el número de estados con que se modela la cadena, para todos los ordenes desde 0 hasta O_{max} ; y se vuelve a repetir el proceso de testear las hipótesis estadísticas mencionadas, partiendo desde el orden 0. Si al reducir el número de estados se llega a uno sólo, dicho orden se descarta en caso de ser mayor a 1; en caso contrario se obtiene la cadena requerida, correspondiente a una de orden 0 y formada por un sólo estado.

La Figura 3.6 ilustra un diagrama de flujo para la metodología descrita.

Los bloques por los que está compuesto este diagrama se detallan a continuación.

- Datos y Parámetros: Corresponde al *input* de la metodología. Está compuesto por el conjunto de datos que se desea modelar mediante una cadena de Markov, y los parámetros de la metodología. Estos parámetros corresponden a los valores de t , p^* , y α_0 (nivel de significancia con que se desea realizar los test de hipótesis).
- Test(n, NE_n): Corresponde a la realización del test de hipótesis para una cadena de Markov de orden n conformada por NE_n estados.
- Test OK: Se verifica si el test de hipótesis falla en rechazar H_0 (Si), o rechaza H_0 (No).
- $\overline{NE} = \overline{NE} - 1$: Se sustrae una unidad al número máximo de estados para todos los órdenes, desde 0 hasta O_{max} .
- Fin: Corresponde al *output* de la metodología. Se obtiene el número de estados y orden de una cadena de Markov homogénea, dados por $NE_{O_{max}}$ y O_{max} respectivamente.

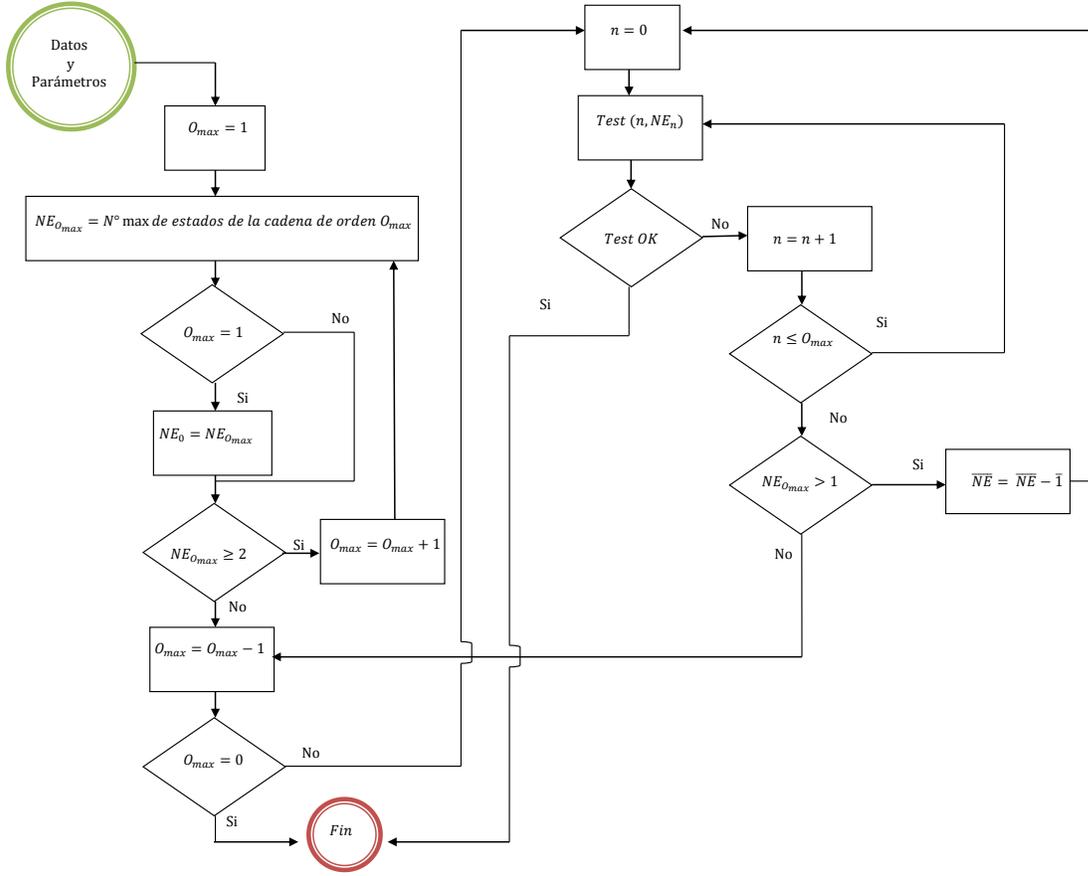


Figura 3.6: Diagrama de flujo que grafica la metodología descrita.

Con el propósito de poder obtener una mejor representación del perfil de corriente mediante la metodología descrita, se ha realizado la siguiente consideración. Se divide este perfil en r intervalos, cuyo valor será explicado más adelante, y se aplica el procedimiento para la determinación del valor de los estados independientemente en cada uno de los intervalos (siempre con el mismo número de centroides para cada intervalo). Luego, una vez obtenida la pertenencia de cada una de las muestras del perfil a su centroide correspondiente, el proceso de la determinación del número de estados y la estimación de las probabilidades de transición se realiza considerando las muestras de todos los intervalos. Para esto se considera de que muestras de diferentes intervalos, pertenecientes a centroides equivalente, corresponden al mismo estado. Es decir, si muestras de diferentes intervalos tienen asociado su centroide correspondiente de mayor valor, entonces se consideran pertenecientes al mismo estado.

Para determinar el valor de los estados de la cadena, se realiza una ponderación exponencial (EWMA, por sus siglas en inglés) [42] de los valores de los estados en cada intervalo. Sea m el número de estados de la cadena, y sea $S(q) = \{S_1(q), \dots, S_m(q)\}$ ($\forall q \in \{1, \dots, r\}$) los valores de dichos estados en el intervalo q . El valor de los estados, para la cadena correspondiente el perfil completo, vienen dados por los valores extremos de la siguiente ecuación recursiva:

$$\tilde{S}_i(q) = (1 - \alpha) \cdot S_i(q) + \alpha \cdot \tilde{S}_i(q - 1) \quad \forall i \in \{1, \dots, m\}, \forall q \in \{2, \dots, r\}. \quad (3.16)$$

En donde, $\tilde{S}_i(q)$ corresponde al valor ponderado del estado i en el intervalo q , cuya condición de borde es $\tilde{S}_i(1) = S_i(1)$. Mientras que α corresponde al factor de olvido de este ponderador.

Mediante esta consideración se busca obtener una mejora en dos aspectos. Por un lado, al realizar el proceso de determinación del valor de los estados separadamente en cada uno de los intervalos, se obtiene una mejor representación de la cadena para cada uno de ellos. Esto se debe a que el valor de los estados en cada intervalo sólo depende de las muestras de dicho intervalo, no viéndose influenciado por el resto de ellos. Así, es posible compensar el hecho de que el ESD; bajo el supuesto que el dispositivo al que energiza funciona con un requerimiento de potencia constante, debe suministrar una mayor corriente a medida que el voltaje en sus bornes decae. Por otro lado, al realizar el proceso de agrupamiento de datos mediante k-means con un conjunto menor de datos, se logra que las soluciones sub-óptimas entregada por este algoritmo se encuentren más cercanas a su óptimo. Esto se refleja en que al desarrollar realizaciones independientes del algoritmo, las soluciones obtenidas presenten menor variabilidad entre sí.

La metodología presentada fue aplicada a los conjuntos de datos Set #1 y Set #2 anteriormente descritos, considerando en ellos sólo los primeros 1800 instantes de tiempo del perfil de corriente. Estos conjuntos fueron segmentados en 6 y 5 intervalos, respectivamente; y se utilizó un factor de olvido de $\alpha = 0.8$ en ambos casos. Los tests estadísticos fueron desarrollados con un nivel de significancia $\alpha_0 = 0.05$, y se usó valores para $t = 0.075$ y $p^* = 0.02$. Sea P la matriz de probabilidades de transición y S un vector conformado por los valores de los estados de la cadena obtenida.

Para el Set #1, se obtuvo una cadena de primer orden conformada por 3 estados. Los parámetros de dicha cadena corresponden a:

$$P = \begin{bmatrix} 0.9656 & 0.0250 & 0.0094 \\ 0.0198 & 0.9703 & 0.0099 \\ 0.0170 & 0.0227 & 0.9602 \end{bmatrix}. \quad (3.17)$$

$$S = (1.9720, 2.2152, 2.5402). \quad (3.18)$$

El valor-p para el test de hipótesis en qué se aceptó los parámetros de esta cadena de Markov corresponde a aproximadamente 0.16.

La Figura 3.7 muestra al perfil de corriente correspondiente a este conjunto de datos, en color negro; y la trayectoria asociada a una cadena de Markov que se obtuvo gracias a la metodología, en color rojo. Se aprecia cómo el valor de los estados de esta trayectoria cambia entre los diferentes 6 intervalos en que se segmentó el conjunto.

Para el Set #2, se obtuvo una cadena de primer orden conformada por 2 estados, cuyos parámetros corresponden a:

$$P = \begin{bmatrix} 0.5709 & 0.4291 \\ 0.5577 & 0.4423 \end{bmatrix}. \quad (3.19)$$

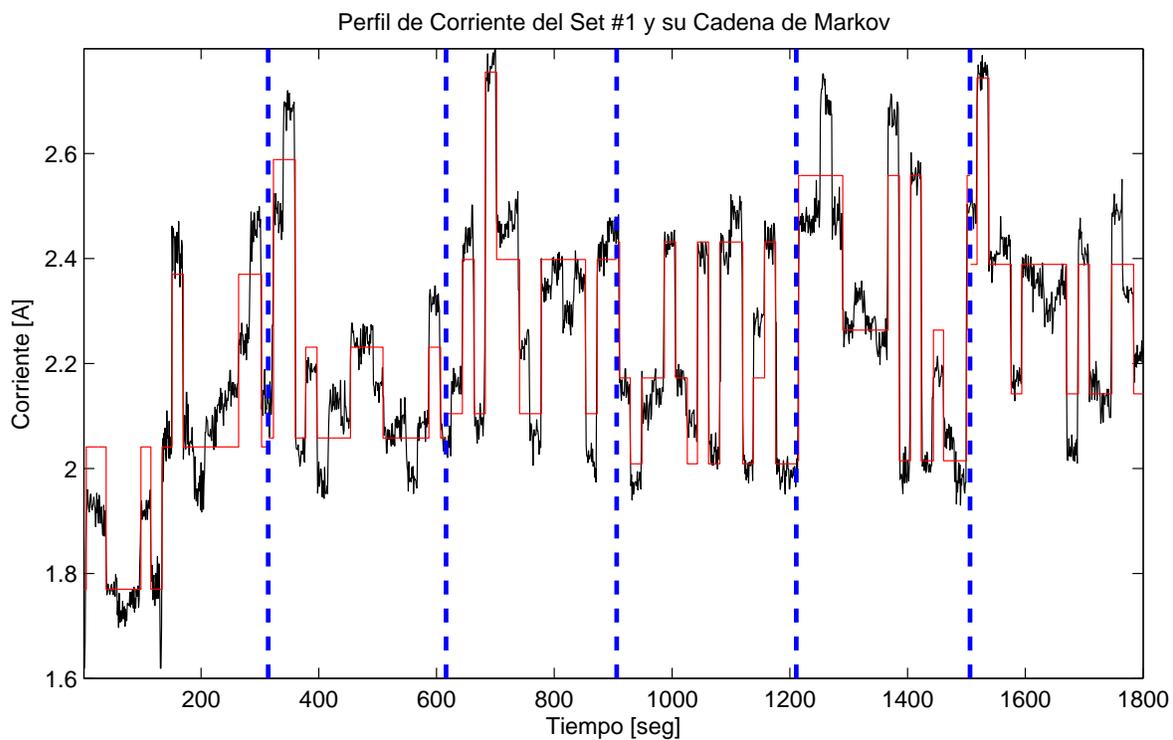


Figura 3.7: Perfil de corriente correspondientes al Set #1 y su trayectoria asociada a una cadena de Markov.

$$S = (1, 3). \tag{3.20}$$

Para este caso, el valor-p corresponde a 0.25.

La Figura 3.8 muestra un acercamiento del perfil de corriente correspondiente a este conjunto, en color negro; y su respectiva trayectoria asociada a una cadena de Markov, mediante puntos de color rojo.

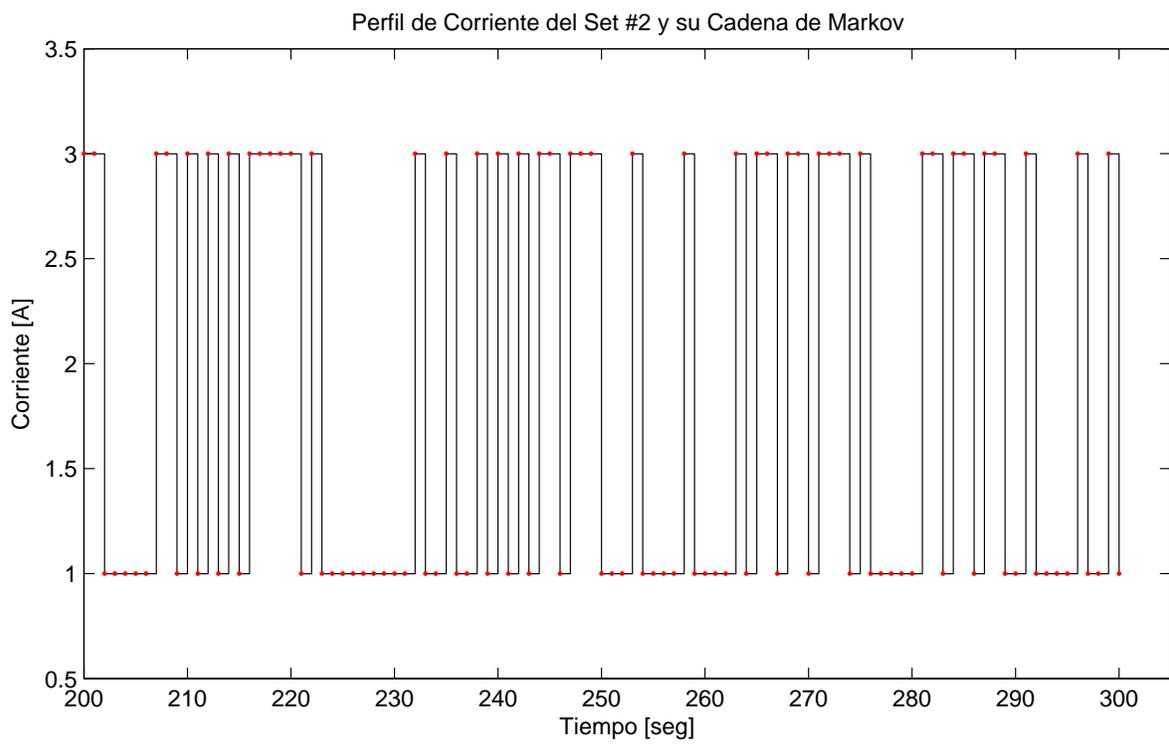


Figura 3.8: Acercamiento del perfil de corriente correspondientes al Set #2 y su trayectoria asociada a una cadena de Markov.

Capítulo 4

Validación de la Metodología

4.1. Modelo de Estado-Espacio para Estimación y Pronóstico del SOC Utilizando Caracterización Estadística de su Perfil de Uso Futuro

El objetivo principal de este trabajo corresponde al desarrollo de una metodología que permita obtener una caracterización estadística del perfil de corriente utilizado en el proceso de descarga de un ESD. El motivo de interés por el que se requiere esta caracterización, es poder desarrollar un esquema de pronóstico en línea para el SOC. Para poder llevar a cabo esto, es necesario disponer de un modelo que represente el proceso de descarga del ESD; el cual debe ser capaz de caracterizar el impacto que producen diferentes perfiles de corrientes de descarga en el voltaje en bornes de la batería, condicional a un valor dado del SOC. Adicionalmente, este modelo debe ser capaz de permitir trabajar en línea; es decir, tanto sus entradas como salidas deben ser variables que puedan ser adquiridas, y posteriormente, procesadas en tiempo real. También el número de dichas variables debe ser acotado, con el objetivo de garantizar un bajo costo computacional. De esta forma, se ha elegido trabajar con los datos de las señales de corriente y voltaje correspondientes al proceso de descarga de una batería.

El modelo que se ha decidido utilizar corresponde a un modelo empírico de espacio-estado inspirado en los circuitos eléctricos equivalentes de la batería. Este modelo, presentado en (4.1), (4.2) y (4.3), intenta explicar la caída de tensión en los bornes de la batería durante su proceso de descarga.

Modelo de Transición de Estados

$$x_1(k+1) = x_1(k) + w_1(k). \quad (4.1)$$

$$x_2(k+1) = x_2(k) - v(k) \cdot i(k) \cdot \Delta t \cdot E_{crit}^{-1} + w_2(k). \quad (4.2)$$

Ecuación de Observación

$$v(k) = V_L + (V_0 - V_L) \cdot e^{\gamma \cdot (x_2(k)-1)} + \alpha \cdot V_L \cdot (x_2(k) - 1) + \dots \quad (4.3)$$

$$\dots(1 - \alpha) \cdot V_L \cdot \left(e^{-\beta} - e^{-\beta \cdot \sqrt{x_2(k)}} \right) - i(k) \cdot x_1(k) + \eta(k).$$

En donde, el perfil de utilización $i(k)$ [A] y el tiempo de muestreo Δt [seg] son variables de entrada, y el voltaje en bornes de la batería $V(k)$ [V] corresponde a la salida del sistema. Los estados son definidos como $x_1(k)$ (parámetro desconocido del modelo) y $x_2(k)$ (SOC, energía remanente de la batería normalizada por E_{crit}), E_{crit} corresponde a la energía total que se espera obtener del ESD (la cual puede ser inferida de su capacidad nominal o curvas de descarga presentes en sus *datasheets*). Para este trabajo se considera que estado x_2 corresponde a una variable adimensional, valor está comprendido entre 1 ó 100 % (batería completamente cargada) y 0 ó 0 % (batería completamente descargada). Los ruidos de proceso ($w_1(k)$ y $w_2(k)$) y observación ($\eta(k)$) son asumidos Gaussianos. Cabe destacar que el ruido de proceso w_2 está correlacionado con el ruido de observación η , ya que la evolución temporal del estado x_2 depende del voltaje medido.

La ecuación de observación representa el comportamiento no lineal típico presente en las curvas OCV vs SOC para baterías de Ion-Litio. Adicionalmente, esta ecuación incorpora el hecho de que al presentarse una corriente a través de la batería, su voltaje en bornes decae debido a la presencia de la impedancia interna propia de la ella. De esta forma, es posible resumir la ecuación de observación puede como:

$$v(k) = v_{oc}(k) - i(k) \cdot |z_p|; \quad (4.4)$$

en donde, $|z_p|$ corresponde al módulo de impedancia interna de la batería. Mientras que $v_{oc}(k)$ es el voltaje en bornes de circuito abierto en el instante k . La curva paramétrica que describe a este voltaje, presente en la ecuación (4.3), ha sido determinada empíricamente.

Adicionalmente a la incorporación de la impedancia interna, otro aspecto fenomenológico que considera el modelo corresponde a la ley de conservación de energía representada en la ecuación (4.2). En esta ecuación se incorpora el hecho de que el SOC del ESD disminuye a medida que se extrae corriente de él.

Los parámetros V_0 , V_L , α , β y γ corresponden a constantes propias de cada batería. Estos valores son calculados *off-line*, mediante un procedimiento cuyo objetivo es disminuir el error cuadrático medio entre la curva de voltaje obtenida al descargar la batería (completamente cargada) a su corriente nominal y la ecuación de observación presentada en (4.3). La Tabla 4.1 presenta los parámetros obtenidos para la baterías asociadas a los conjuntos de datos presentados en el Capítulo 3.

El valor z_p presentado en la Tabla 4.1 corresponde a una aproximación del valor absoluto de la impedancia interna de la batería. Para determinar dicha aproximación; en la prueba de descarga a corriente constante, realizada para determinar los parámetros de la ecuación de observación, se

Batería	α	β	γ	V_0	V_L	E_{crit}	Z_p
Set #1	0.0776	16	19.6445	4.12	3.9773	20127	0.3
Set #2	0.1500	12	6.6061	4.00	3.8126	19865	0.2

Tabla 4.1: Parámetros del modelo para las baterías asociadas a los conjuntos de datos Set #1 y Set #2.

agregan dos pulso en la corriente de descarga. De esta forma, es posible estimar el valor de dicho parámetro mediante $|z_p| = |\frac{\Delta V}{\Delta I}|$. En donde, ΔI representa la diferencia de corriente empleada para realizar el pulso en la prueba. Mientras que ΔV corresponde a la diferencia de voltaje en bornes de la batería producida al aplicar dicho pulso de corriente.

Esta prueba de descarga también es utilizada para obtener una estimación *a priori* del comportamiento del SOC, mediante la utilización de la ley de conservación de energía. Esta estimación es utilizada para calcular el ancho de los intervalos temporales, con los que se segmenta el perfil de corriente, mencionados en el Capítulo 3. En este trabajo se ha escogido un ancho para cada intervalo dado por el ancho de los intervalos donde el SOC disminuye 10 % en la estimación mencionada.

El modelo propuesto tiene la capacidad de realizar la estimación de parámetros cuyo valor exacto es desconocido. Este es el caso del valor absoluto instantáneo de la impedancia interna de la batería, representada por el estado x_1 . La ecuación (4.1) corresponde a la realización de un proceso de evolución artificial [43], con el objetivo de realizar una correcta estimación de este estado. La realización de este proceso ayuda a compensar errores debidos a una inadecuada elección de la condición inicial para el estado, dada por z_p . Por otro lado, esta estrategia de estimación tiene la ventaja de poder incorporar la incidencia de factores externos, cuyo efecto no está contemplado en la formulación del modelo de espacio-estado. Un ejemplo de tales factores es la temperatura ambiental, ya que el valor de la impedancia interna se ve afectada por ella.

La varianza del ruido de observación η se determinó mediante un procesamiento estadístico *off-line*, el cual consiste en aplicar un filtro pasa bajo a la señal de voltaje medido, y posteriormente calcular el valor absoluto de la diferencia entre la señal original y la filtrada. La varianza empírica de dicha diferencia se utilizó como la varianza de este ruido, ya corresponde a un estimador de ella. Por otra parte, la varianza de w_2 fue determinada mediante ajustes empíricos a través de un análisis de sensibilidad realizada a la adaptabilidad del filtro en la etapa de estimación, incorporando el efecto del ruido de observación. De forma análoga se determinó la varianza de w_1 .

Dada la no linealidad del modelo de espacio-estado presentado, no se cuenta con una herramienta que permita calcular la solución al problema de filtrado óptimo. Es por esto que se considera la utilización de una solución sub-óptima, la cual permita trabajar en línea. Específicamente se decide trabajar con el algoritmo de Filtro de Partículas, tanto para desarrollar la etapa de estimación y su posterior etapa de pronóstico. Esta herramienta posee la capacidad de ser utilizada en sistemas no lineales y no Gaussianos, siendo atractiva para el modelo fenomenológico utilizado.

Al comienzo de la etapa de estimación, es común que no se tenga conocimiento certero sobre la cantidad de energía almacenada en la batería. Esto se traduce en un desconocimiento de la condición inicial del SOC, y por lo tanto del estado x_2 . Es crítico poder corregir errores que se generan de asumir una cierta condición inicial para dicho estado; ya que para realizar una adecuada etapa de pronóstico del SOC, se requiere que la etapa previa de estimación para dicha variable haya sido desarrollada correctamente. Es por esto que se ha adoptado en este trabajo una estrategia de aprendizaje adaptativo, o *adaptive learning strategy*, [44] para dicho estado; la cual es factible de ser realizada dentro de un módulo de estimación basado en Filtro de Partículas. Esta estrategia consiste en aumentar la incertidumbre asociada al estado x_2 en (4.2) (representada por la varianza del ruido w_2), asumiendo que cambios en el valor de la impedancia interna, y por lo tanto en el estado x_1 , son despreciables para una ciclo específico de descarga [45]. A pesar de que esta estrategia ayuda a reducir el error asociado a la condición inicial errónea de este estado, y por lo tanto a su convergencia; incorpora fuentes de incertidumbre adicionales, la cual son añadidas de forma artificial, al módulo de estimación si se mantienen invariantes en el tiempo. Es por esto que, una vez que se han adquirido una pequeña cantidad de mediciones secuenciales de corriente y voltaje, la varianza del ruido es reducida exponencialmente hasta alcanzar cierta cota inferior (la cual corresponde a la varianza para w_2 explicada anteriormente en este capítulo).

Una vez realizada la etapa de estimación de los estados, se utiliza dicha información como condiciones iniciales de la etapa de pronóstico. Para esto, se consideró un modulo de pronóstico basado en Filtro de Partículas, el cual hace uso de la distribución *a priori* de la predicción de la trayectoria de los estados a lo largo del horizonte de predicción, junto con una caracterización estadística del perfil de utilización de la batería; con el objetivo de caracterizar la incertidumbre del perfil de descarga futuro de la batería. Esta caracterización estadística corresponde a representar dichos perfiles de corriente como una cadena de Markov homogénea, tal como se describió en el Capítulo 3. Una de las principales ventajas de la utilización de este enfoque para la etapa de pronóstico es que permite realizar la estimación de la pdf del EOD, mediante (2.66), en tiempo real.

4.2. Medidas de Desempeño para Módulos de Pronóstico Basados en Filtro de Partículas

Una vez estimada la distribución del EOD, es posible obtener medidas de desempeño para la etapa de predicción. Dentro de estas medidas, la más común corresponde al valor esperado del EOD, el cual corresponde al valor medio del instante de tiempo en donde ocurre el evento crítico, y se presenta en la ecuación (4.5). Sin embargo, esta métrica no incorpora el concepto de riesgo, ya que considera una gran probabilidad de que el evento crítico se presente antes de dicho valor. Para incorporar este concepto se incluyen otras medidas de desempeño, tales como intervalos de confianza, varianza del EOD y $JIPT_{\gamma\%}$, (Just-in-Time Point, por sus siglas en inglés) [46]. Esta última métrica, presentada en la ecuación (4.6), corresponde al instante de tiempo en que la probabilidad de falla alcanza un umbral $\gamma\%$ predefinido.

$$\widehat{EOD} = \mathbb{E}\{EOD\} = \mathbb{E}\{k \mid \mathbb{E}\{x_2(k)\} = 0\}. \quad (4.5)$$

$$JITP_{\gamma\%} = \arg \min_{eod} \{\mathbb{P}(EOD \leq eod) \geq \gamma\% \}. \quad (4.6)$$

4.3. Determinación de los Parámetros del Algoritmo de Filtro de Partículas Utilizado

Diversos estudios han utilizado algoritmos de pronóstico basados en Filtro de Partículas ([5], [7], [15], [21], [44], [47]). Sin embargo, los parámetros del algoritmo deben ser fijados dependiendo del problema en específico en el que se desea trabajar. Estos parámetros dependen del grado de no linealidad presentado por el modelo, su número de estados, entre otros. Para el caso de estudio de este trabajo, para realizar una correcta estimación de la pdf del EOD es crítico determinar los siguientes parámetros del algoritmo de pronóstico: (i) el número de partículas que son necesarias para representar la pdf de los estados en cada iteración del modelo predictivo estocástico, (ii) el número de iteraciones del algoritmo de pronóstico necesarias para asegurar estándares determinados a cuanto a la exactitud de la predicción de la pdf del EOD, y (iii) el número de realizaciones de la cadena de Markov necesarias para caracterizar adecuadamente los diversos perfiles de utilización futura a los que puede ser sometido el ESD, en donde cada una de estas realizaciones corresponde a un posible perfil de uso representado por una secuencia de estados del valor de la corriente.

En [48] se realizó un estudio para determinar el valor apropiado de estos parámetros para el caso de la estimación de la pdf del EOD de una batería de Ion-Litio. En específico, se realizaron diversos experimentos basados en Filtro de Partículas para el caso de un modelo simplificado, los cuales consistían en hacer variar tanto el número de partículas como el número de realizaciones de la cadena de Markov que son necesarias para caracterizar de forma adecuada la incertidumbre del perfil de descarga futuro de la batería. El modelo simplificado utilizado corresponde a un modelo en espacio-estado lineal y Gaussiano, por lo que la solución óptima para la estimación de la pdf del EOD es posible de obtener analíticamente. El resultado de los experimentos fue comparado con esta solución analítica, mediante las medidas de desempeño (4.5) y (4.6).

El $JITP_{\gamma\%}$ es crítico para la determinación del número de partículas que son necesarias para la representación de la incertidumbre del sistema, ya que entrega información sobre las colas de las distribuciones. También, como se explica en [48], la precisión de la esperanza del EOD depende fuertemente, tanto del número de realizaciones del algoritmo de pronóstico, como del número de realizaciones de la cadena de Markov.

En específico, mediante un análisis de los resultados obtenidos mediante el Filtro de Partículas y la solución óptima analítica, en [48] se concluyó que un número de 40 partículas es apropiado para la implementación del algoritmo, teniendo en cuenta que también se considera el costo computacional asociado la dicha implementación. En cuanto al número de realizaciones del algoritmo de pronóstico, un número de 25 es suficiente para obtener estimaciones adecuadas para el $JITP_{5\%}$, $JITP_{10\%}$ y $JITP_{50\%}$.

Cabe mencionar que debido al alto costo computacional de realizar distintas iteraciones del algoritmo de pronóstico y realizaciones de la cadena de Markov e inspirado en el trabajo realizado en [22], se consideró la simplificación de considerar sólo una iteración del algoritmo para cada una de dichas realizaciones. Así al menos 25 de estas realizaciones deben ser consideradas. Sin embargo, sólo fueron necesarias 25 realizaciones de la cadena de Markov para proveer una confiable estimación del EOD.

4.4. Resultados del Módulo de Pronóstico Basado en Filtro de Partículas y Caracterización Estadística del Perfil de Uso Futuro

Con el objetivo de validar la metodología presentada para obtener una caracterización estadística del perfil de uso de una batería, se utiliza el modelo en espacio-estado descrito por las ecuaciones (4.1), (4.2) y (4.3) en un módulo de pronóstico basado en Filtro de Partículas. Este módulo de pronóstico considera dicha caracterización estadística para el perfil de corriente de descarga futuro de la batería. Para evaluar los resultados obtenidos de la etapa de pronóstico, se utilizan las medidas de desempeño (4.5) y (4.6).

Los parámetros del algoritmo empleado; número de partículas, iteraciones de la etapa de predicción y realizaciones de la cadena de Markov, corresponden a los descritos en la sección anterior. Es decir; 40 partículas, 25 iteraciones y 25 realizaciones. Mientras que los datos de validación corresponden a los conjuntos Set #1 y Set #2.

El instante de tiempo real en que ocurre la descarga completa de la batería, también conocido como *ground truth*, para los conjuntos Set #1 y Set #2 corresponde a 2738 [seg] y 3381 [seg] respectivamente. Mientras que la condición inicial para el estado x_2 fue elegida arbitrariamente mediante la generación de una variable aleatoria en el intervalo $[0.8; 0.9]$ para ambos conjuntos, a pesar del conocimiento previo que se tenía que ambas pruebas de descarga fueron realizadas con la batería completamente cargada. El motivo de esta elección aleatoria para el SOC inicial de la batería corresponde a poder observar cómo es el desempeño de la etapa de filtrado cuando se tiene incertidumbre en la condición de dicho estado.

Dada la naturaleza estocástica del Filtro de Partículas, la validación de la metodología propuesta se lleva a cabo mediante varias realizaciones, aleatorias e independientes, del algoritmo. Cada realización consiste en una iteración del algoritmo de estimación y 25 iteraciones del módulo de pronóstico (cada una asociada a una realización independiente de cadena de Markov). Para ambos conjuntos de datos, la etapa de pronóstico se desarrolla una vez transcurridas las primeras 1800 muestras de datos; y sus condiciones iniciales vienen dadas por el resultado del proceso de estimación. Las Figuras 4.1, 4.2 y 4.3 corresponden sólo a una de dichas realizaciones para el

conjunto de datos Set #1. De igual forma, las Figuras 4.4, 4.5 y 4.6 ilustran una realización en particular para el conjunto Set #2. Mientras que las Tablas 4.2 y 4.3 muestran los resultados de todas las realizaciones hechas para los conjuntos Set #1 y Set #2 respectivamente.

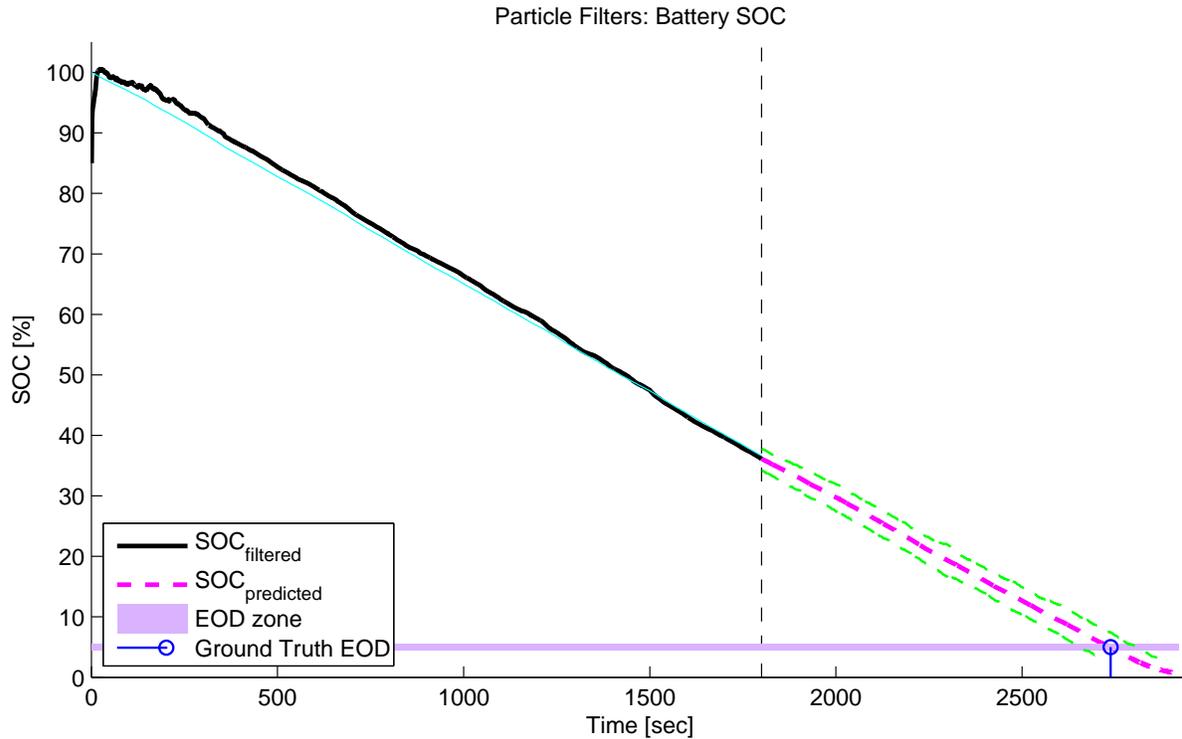


Figura 4.1: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): SOC estimado (línea negra), SOC pronosticado (línea punteada magenta), e intervalo de confianza del 95 % para la predicción del SOC (línea punteada verde).

Nº	$E\{EOD\}$ [seg]	Intervalo de Confianza - 95 % [seg]	JITP _{5%} [seg]	JITP _{15%} [seg]
1	2574	[2447; 2736]	2474	2502
2	2741	[2708; 2773]	2705	2717
3	2690	[2480; 2900]	2498	2556
4	2704	[2619; 2788]	2631	2654
5	2733	[2630; 2838]	2638	2672

Tabla 4.2: Resultados de diferentes realizaciones del módulo de pronóstico para el SOC (Conjunto de datos Set #1). Ground Truth del EOD: 2738 [seg].

Es posible apreciar en las Figuras 4.1 y 4.4 que el proceso de estimación realizado para el estado x_2 es capaz de corregir errores asociados a una condición inicial errónea mediante la aplicación de la estrategia de aprendizaje adaptativa anteriormente descrita. Esta corrección es hecha de forma rápida, ya que transcurridas pocos instantes de tiempo, el valor estimado de x_2 converge a su valor real. Esto muestra que el módulo de estimación propuesto es capaz de entregar estimaciones confiables para el SOC. Por otro lado, en las Figuras 4.2 y 4.5 se observa que la curva de voltaje pronosticado efectivamente incorpora la caída abrupta presente en el tipo de ESD utilizados cuando el SOC cae por debajo del 10%. Se aprecia que en esta zona del proceso de descarga hay un sesgo entre el voltaje observado y el pronosticado. Sin embargo, este hecho no debería ser

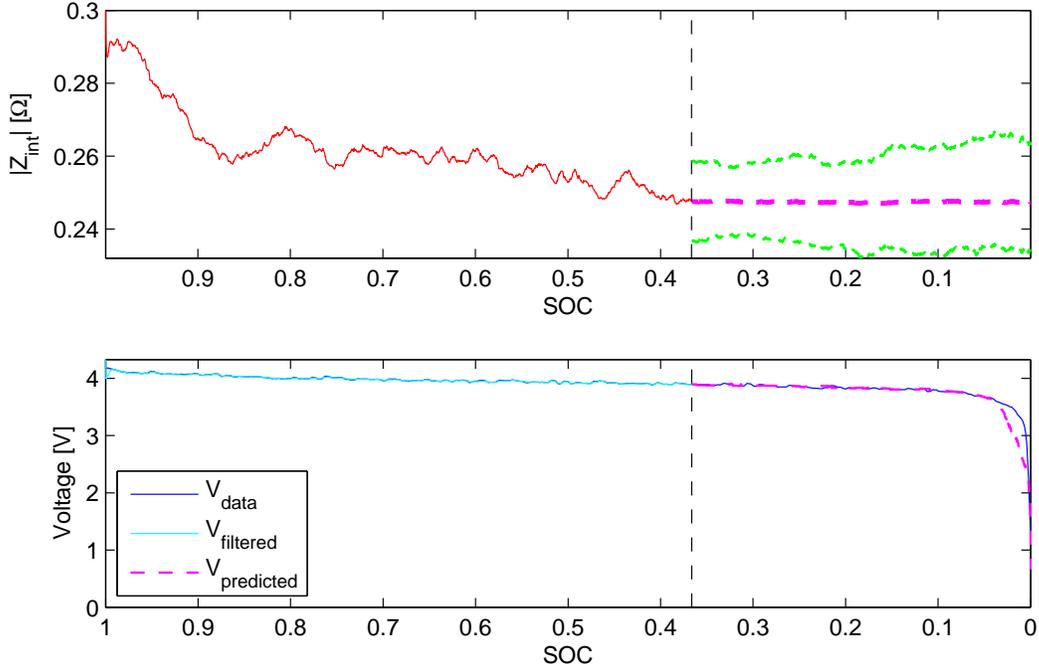


Figura 4.2: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): a) Evolución del estado x_1 , en función del SOC, durante la etapa de estimación (línea roja) y predicción (línea punteada magenta). b) Voltaje medido (línea azul), voltaje estimado (línea cyan), y voltaje predicho (línea punteada magenta).

N°	$\mathbb{E}\{EOD\}$ [seg]	Intervalo de Confianza - 95 % [seg]	JITP _{5%} [seg]	JITP _{15%} [seg]
1	3289	[3233; 3345]	3237	3256
2	3311	[3262; 3360]	3260	3279
3	3284	[3242; 3326]	3237	3255
4	3343	[3252; 3434]	3259	3289
5	3336	[3283; 3388]	3278	3300

Tabla 4.3: Resultados de diferentes realizaciones del módulo de pronóstico para el SOC (Conjunto de datos Set #2). Ground Truth del EOD: 3381 [seg].

preocupante para el caso de estudio presentado en este trabajo. Esto se debe a que en aplicaciones reales la batería no es descargada hasta tales extremos (debido a especificaciones del fabricante, requerimientos de potencia del equipo energizado, protocolos de seguridad, entre otros), sino que se utiliza sólo hasta el principio de la zona donde se origina dicha caída abrupta.

La Figura 4.7 ilustra el perfil de corriente correspondiente al Set #1 consumido durante la etapa de estimación, en color negro; y la trayectoria asociada a una cadena de Markov utilizada para la caracterización de dicho perfil, en color rojo. Las líneas punteadas en color azul indican los intervalos en que se segmentó la etapa de estimación, mientras que la línea punteada verde indica el instante en que se realiza la etapa de pronóstico. Posterior a dicho instante, la señal en color rojo ilustra una de las realizaciones de la cadena de Markov utilizada en esta etapa y obtenida con

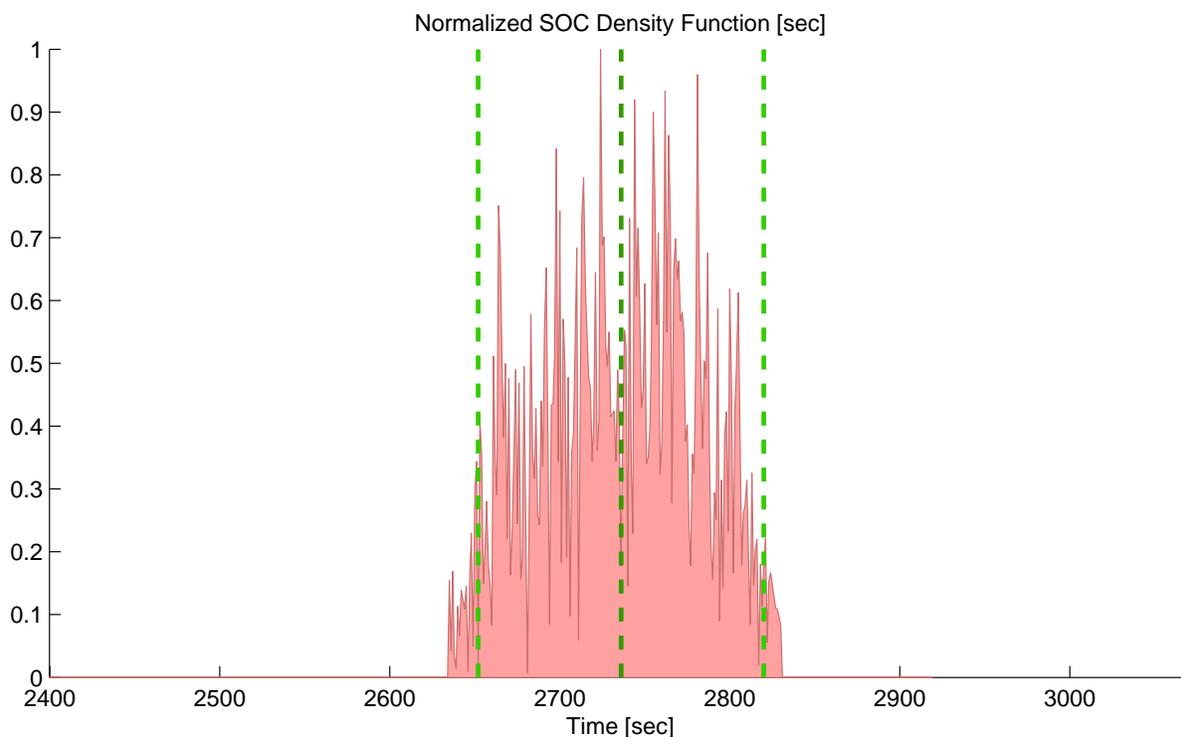


Figura 4.3: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #1): Estimación de la pdf del EOD (normalizada). Líneas verticales punteadas indican los límites del intervalo de confianza del 95 % (líneas exteriores) y la esperanza (línea interior) para esta pdf.

la metodología propuesta.

Por otro lado, la Figura 4.8 ilustra un acercamiento del perfil de corriente correspondiente al Set #2 consumido durante la etapa de estimación, en color negro; y la trayectoria asociada a una cadena de Markov utilizada para la caracterización de dicho perfil, mediante puntos de color rojo. La línea punteada verde indica el instante en que se realiza la etapa de pronóstico. Posterior a dicho instante, la señal en color rojo ilustra una de las realizaciones de la cadena de Markov utilizada en esta etapa y obtenida con la metodología propuesta.

Mediante los resultados presentados en las Tablas 4.2 y 4.2 es posible analizar cuantitativamente el desempeño del módulo de pronóstico para el SOC de un ESD, el cual utiliza la metodología para la caracterización del perfil de uso futuro propuesta en este trabajo.

Es posible apreciar en la Tabla 4.2 que el valor esperado del EOD varía en cada realización del algoritmo, por lo que corresponde a una variable aleatoria. Más aún, los resultados obtenidos para las realizaciones de dicha variable aleatoria muestran que su valor puede subestimar o sobreestimar el valor del ground truth. Sin embargo, se observa que los valores obtenidos tienden a subestimar dicho valor; siendo el mayor error de sobrestimación obtenido de 3 [seg], en una horizonte de predicción de 938 [seg]. Bajo el supuesto de una enfoque conservador, esto corresponde a un resultado deseable; ya que en aplicaciones reales es mucho más crítico la obtención de una sobrestimación. Esto se debe a que la toma de decisiones basada en dicha sobrestimación puede llevar a la planificación de una misión de uso del equipo conectado al ESD que es imposible de

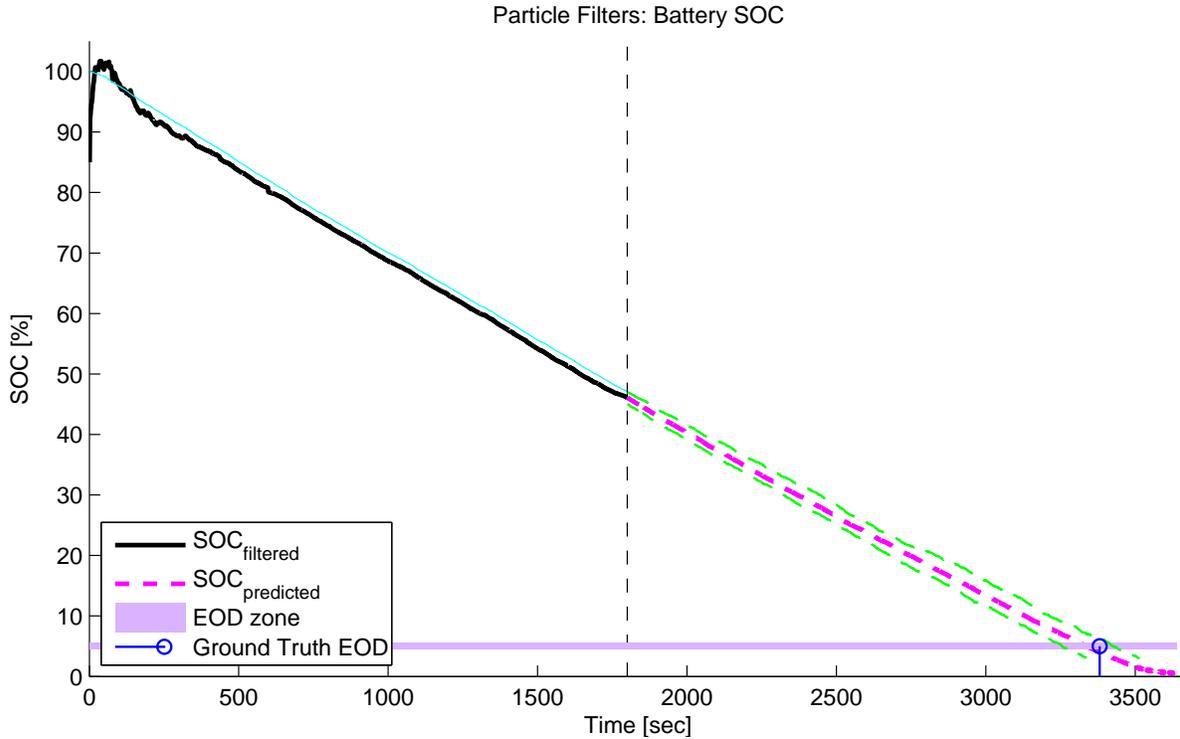


Figura 4.4: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): SOC estimado (línea negra), SOC pronosticado (línea punteada magenta), e intervalo de confianza del 95 % para la predicción del SOC (línea punteada verde).

llevar a cabo. Por otro lado, todos los valores obtenidos para el $JITP_{5\%}$ y $JITP_{15\%}$ son menores al ground truth, por lo que una toma de decisiones basada en este indicador garantiza una utilización seguro del ESD. La máxima diferencia entre el ground truth y los valores del $JITP_{5\%}$ - $JITP_{15\%}$ corresponden a 264 - 236 [seg] respectivamente (representando un 28.1 - 25.2 % del horizonte de predicción). La Figura 4.3 ilustra la pdf estimada del EOD para una de las realizaciones del módulo de pronóstico.

De igual forma, los resultados presentados en la Tabla 4.2 indican que el valor esperado del EOD corresponde a una variable aleatoria. Sin embargo, los resultados obtenidos para las realizaciones de dicha variable muestran que su valor siempre subestima el valor del ground truth. El mayor valor obtenido para dicha subestimación corresponde a 97 [seg], en un horizonte de predicción de 1581 [seg]. Como fue comentado anteriormente, esta subestimación es deseable desde el punto de vista de un enfoque conservador, ya impide una decisión errónea con respecto a la autonomía del dispositivo. Al igual que el caso anterior, los resultados de esta tabla indican que los valores para el $JITP_{5\%}$ y $JITP_{15\%}$ son siempre menores al valor del ground truth. La máxima diferencia entre el ground truth y los valores del $JITP_{5\%}$ - $JITP_{15\%}$ corresponden a 144 - 126 [seg] respectivamente (representando un 9.1 - 8.0 % del horizonte de predicción). La Figura 4.6 ilustra la pdf estimada del EOD para una de las realizaciones del módulo de pronóstico.

Así, tomar una decisión respecto a la autonomía del ESD, basada en los indicadores del $JITP_{\gamma\%}$, garantiza que no se sobrestime dicha autonomía y que el error con respecto al ground truth está

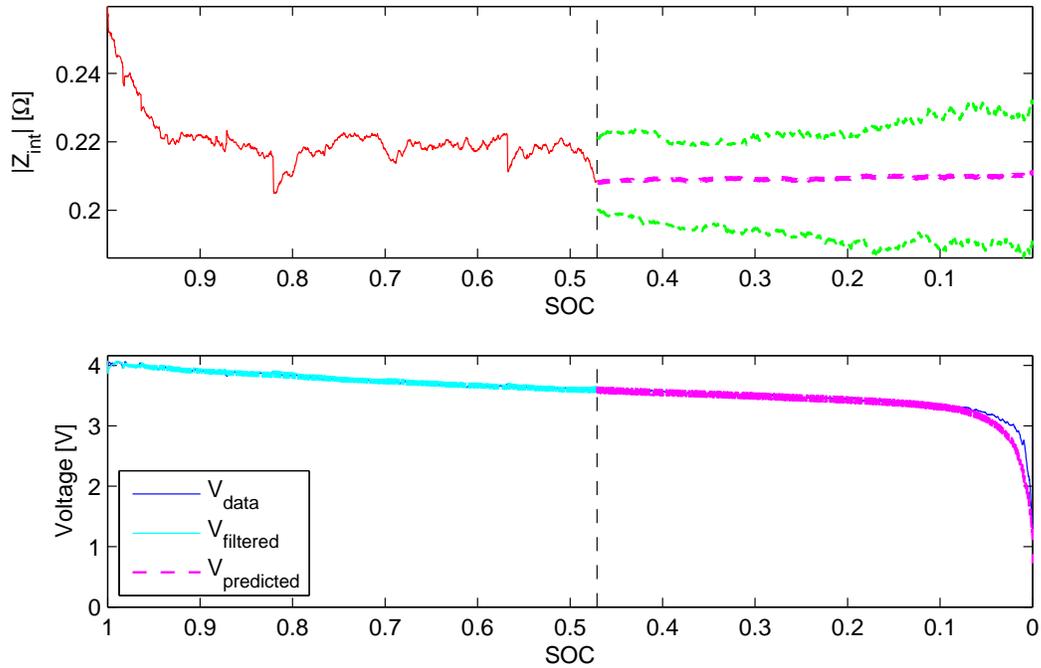


Figura 4.5: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): a) Evolución del estado x_1 , en función del SOC, durante la etapa de estimación (línea roja) y predicción (línea punteada magenta). b) Voltaje medido (línea azul), voltaje estimado (línea cyan), y voltaje predicho (línea punteada magenta).

acotado en un porcentaje del horizonte de predicción.

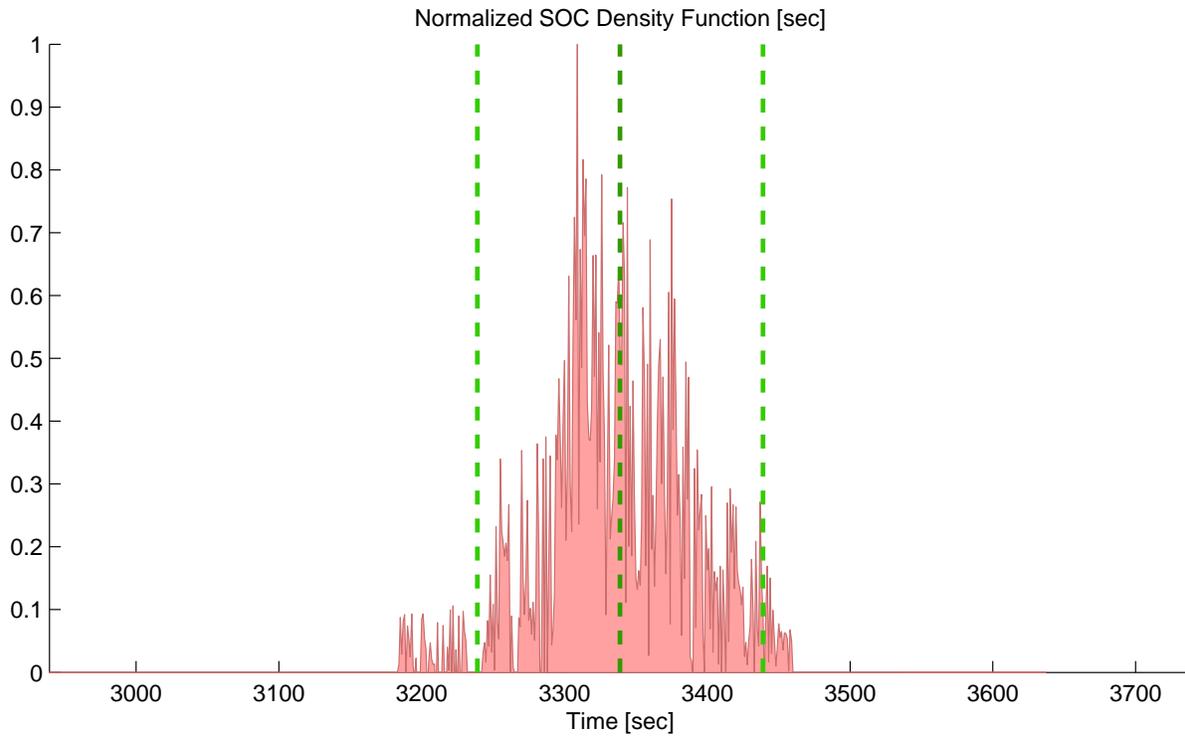


Figura 4.6: Pronóstico del SOC utilizando un módulo basado en Filtro de Partículas (Conjunto de datos Set #2): Estimación de la pdf del EOD (normalizada). Líneas verticales punteadas indican los límites del intervalo de confianza del 95 % (líneas exteriores) y la esperanza (línea interior) para esta pdf.

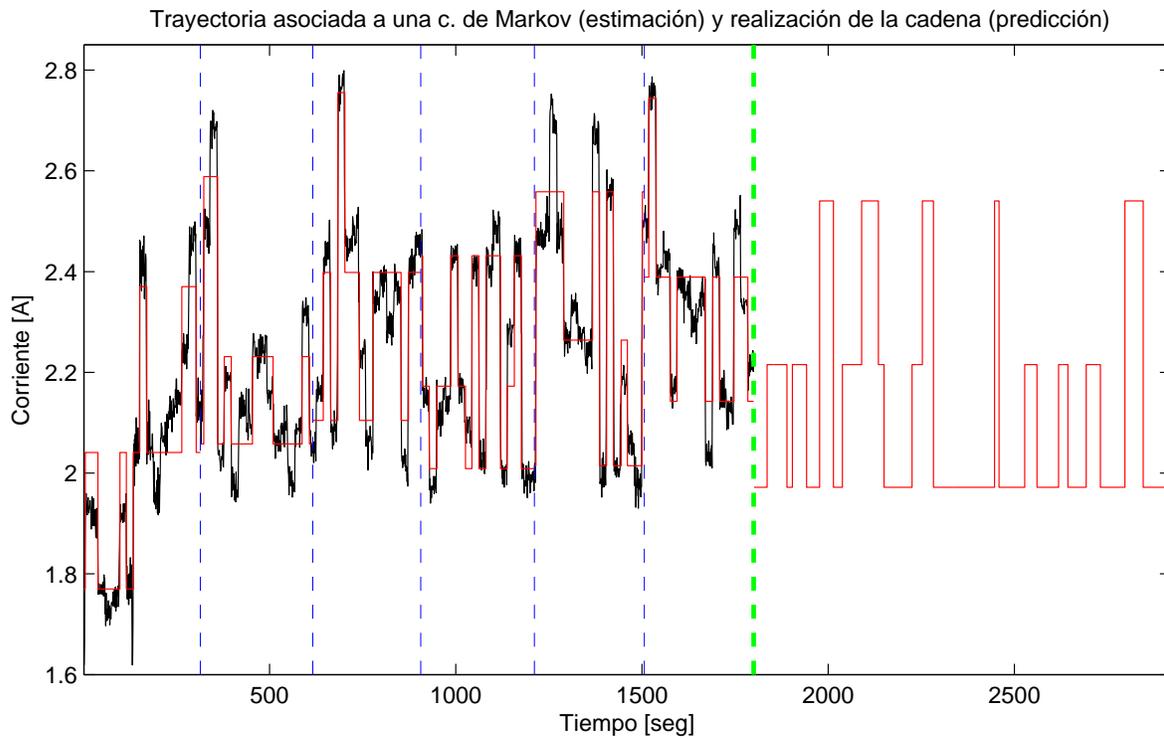


Figura 4.7: Perfil de corriente correspondiente al Set #1 y la trayectoria asociada a una cadena de Markov (etapa de estimación); junto con una realización de la cadena de Markov (etapa de pronóstico).

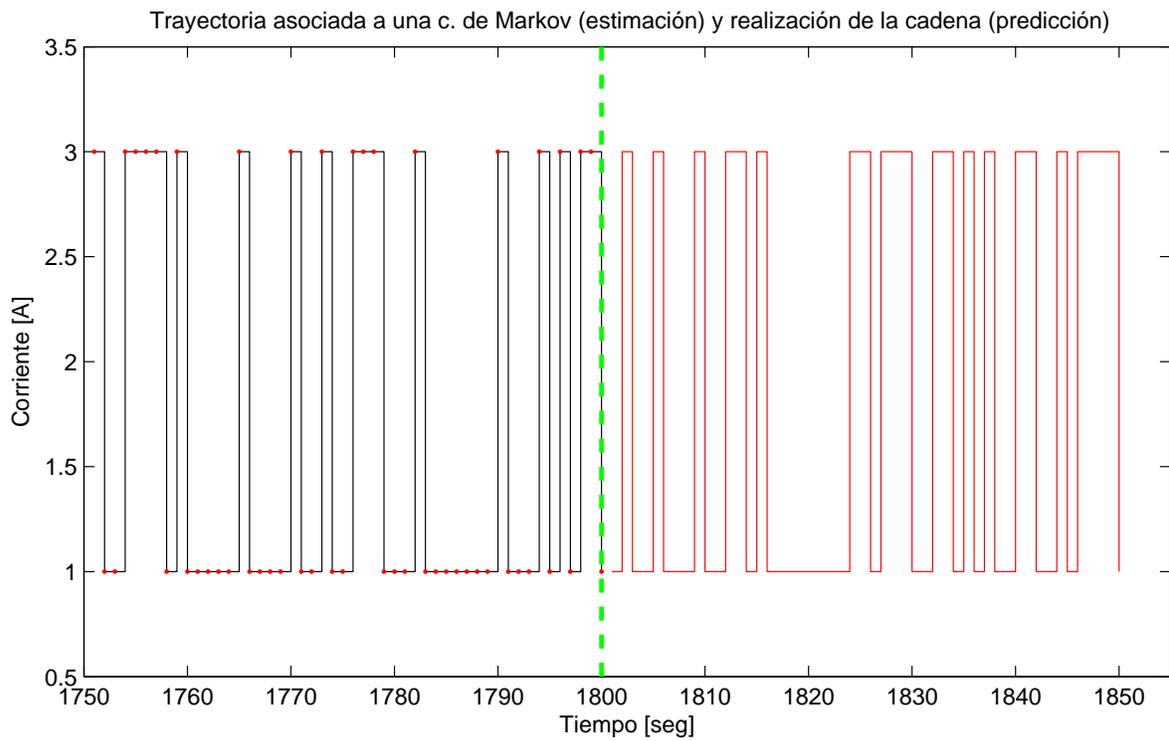


Figura 4.8: Acercamiento del perfil de corriente correspondiente al Set #2 y la trayectoria asociada a una cadena de Markov (etapa de estimación); junto con una realización de la cadena de Markov (etapa de pronóstico).

Capítulo 5

Conclusiones

En este trabajo se presenta una metodología para realizar una caracterización estadística del perfil de uso de una batería. Esta metodología permite obtener los mejores parámetros necesarios para modelar el perfil de descarga (señal de corriente) de una ESD mediante una cadena de Markov homogénea. Estos parámetros corresponden al valor de los estados, las probabilidades de transición y el orden de la cadena. La determinación de cada uno de ellos es justificada mediante herramientas matemáticas, por lo que se dejan de lado los enfoques heurísticos; los cuales si bien pueden dar buenos resultados en la práctica, carecen de un fundamento teórico que garantice un buen comportamiento. Los fundamentos con que se diseñó la metodología corresponden a herramientas clásicas de teoría de probabilidades y estadística, por lo que no presenta inconvenientes de una complejidad teórica sofisticada. Más aun, la metodología es lo suficientemente general para poder ser utilizada en muchas áreas de la ingeniería, no restringiéndose sólo a su uso para la caracterización de perfiles de corriente.

Esta metodología corresponde a una solución novedosa al problema de caracterización estadística de los perfiles de uso, no habiéndose encontrado trabajos similares durante el proceso de revisión bibliográfica. Esto conlleva a que lo presentado en este trabajo pueda contribuir a enriquecer los conocimientos de esta disciplina, llegando a ser un aporte al estado del arte. Cabe mencionar que la solución entregada por esta metodología en caso que se rechacen todos los test de hipótesis utilizados, entrega la solución clásica para el problema de caracterización (valor medio del perfil). En este sentido, lo propuesto aquí puede ser considerado una generalización para dicha solución clásica, y no un enfoque alternativo o desligado de ella.

Dado que el objetivo final de realizar dicha caracterización es realizar su aplicación al pronóstico del Estado-de-Carga y estimación de la pdf del EOD, en este trabajo se presentó un módulo de estimación y pronóstico para este parámetro basado en Filtro de Partículas. Este módulo utiliza en la etapa de pronóstico la metodología presentada, con el fin de poder evaluar su desempeño y poder validarla.

Dos conjuntos de datos, correspondientes al proceso de descarga de dos baterías diferentes,

fueron utilizados para la validación. El primero de estos conjuntos corresponde a un perfil de descarga consistente en cambios aleatorios de diversos puntos de operación. Mientras que el segundo corresponde a la realización de una cadena de Markov homogénea de dos estados. Al aplicar la metodología para estos dos conjuntos, se logró obtener una cadena homogénea de primer orden que los representara. Más aun, para el segundo caso la cadena obtenida corresponde exactamente al mismo perfil de corriente utilizado. Esto muestra que la metodología cumple su objetivo de obtener los parámetros adecuados de una cadena de Markov que representa un perfil de uso.

Para la evaluación de los resultados obtenidos en la etapa de pronóstico se utilizaron medidas de desempeño *ad-hoc*. Éstas indicaron, para ambos conjuntos de datos de validación, que mediante la utilización de la metodología presentada, es posible obtener resultados aceptables en términos de exactitud y precisión para el pronóstico del SOC y estimación de la pdf del EOD. Pudiendo de esta forma ser factible de ser incorporada dentro de un BMS para aplicaciones reales.

Como trabajo futuro se plantea realizar un análisis comparativo entre esta metodología y los métodos presentes en la literatura, comparando los resultados obtenidos en la etapa de pronóstico y los recursos computacionales utilizados para la obtención de la caracterización estadística. Por otro lado, si bien la metodología propuesta tiene consistencia teórica, su validación sólo se realizó con dos conjuntos de datos. Es por esto que se plantea también estudiar su desempeño con un gama más amplia de ellos, no necesariamente correspondientes a descargas de baterías de Ion-Litio.

Bibliografía

- [1] C. Chen W. He, N. Williard and M Pecht. State of charge estimation for electric vehicle batteries using unscented kalman filtering. *Microelectronics Reliability*, 53:840–847, 2013.
- [2] The Consumer Electronics Association. *Electric vehicles: the future of driving*, 2010.
- [3] C. Sankavaram B. Pattipati and K. Pattipati. System identification and estimation framework for pivotal automotive battery management system characteristics. *Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, 41:869–884, 2011.
- [4] A. Khoobroo A. Ranjbar, A. Banaei and B. Fahimi. Online estimation of state of charge in li-ion batteries using impulse response concept. *Smart Grid, IEEE Transactions on*, 3(1):360–367, March 2012.
- [5] B. Saha and K. Goebel. Modeling li-ion battery capacity depletion in a particle filtering framework. In *Annual Conference of the Prognostics and Health Management Society*, San Diego, CA, 2009.
- [6] P. Singh T. Atwater A. J. Salkind, C. Fennie and D. Reisner. Determination of state-of-charge and state-of-health of batteries by fuzzy logic methodology. *Journal of Power Sources*, 80:293–300, 1999.
- [7] M. Orchard and G. Vachtsevanos. A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control 2009*, 31:221–246, 2009.
- [8] D. Petreus D. Cadar and C. Orian. A method of determining a lithium-ion battery s state of charge. In *15th International Symposium for Design and Technology of Electronics Packages (SIITME) 2009*, pages 257–260, September 2009.
- [9] C. Naxin S. Qingsheng, Z. Chenghui and Z. Xiaoping. Battery state-of-charge estimation in electric vehicle using elman neural network method. In *29th Chinese Control Conference (CCC) 2010*, pages 5999–6003, July 2010.
- [10] M. Yan Z. Di and B. Qing-Wen. Estimation of lithium-ion battery state of charge. In *30th Chinese Control Conference*, 2011.
- [11] J. Lin X. Tang, X. Mao and B. Koch. Li-ion battery parameter estimation for state of charge. In *American Control Conference (ACC) 2011*, pages 941–946, June-July 2011.

- [12] B. Cao J. Xu, C. Mi and J. Cao. A new method to estimate the state of charge of lithium-ion batteries based on the battery impedance model. *Journal of Power Sources*, 233:277–284, July 2013.
- [13] M. Charkhgard and M. Farrokhi. State-of-charge estimation for lithium-ion batteries using neural networks and ekf. *IEEE Transactions on Industrial Electronics*, 57(12):4178–4187, December 2010.
- [14] W. Haiying L. Ran, W. Junfeng and L. Gechen. Prediction of state of charge of lithium-ion rechargeable battery with electrochemical impedance spectroscopy theory. In *Industrial Electronics and Applications (ICIEA) 2010 the 5th IEEE Conference on*, pages 684–688, June 2010.
- [15] J. Ma M. Dalal and D. He. Lithium-ion battery life prognostic health management system using particle filtering framework. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 225:81–90, 2011.
- [16] B. Blunier N. Watrin and A. Miraoui. Review of adaptive systems for lithium batteries state-of-charge and state-of-health estimation. In *Transportation Electrification Conference and Expo (ITEC), 2012 IEEE, 2012*.
- [17] Y. Hu and S. Yurkovich. Battery cell state-of-charge estimation using linear parameter varying system techniques. *Journal of Power Sources* 198, pages 338–350, 2012.
- [18] C. Blanco Viejo J. Alvarez Anton, P. Garcia Nieto and J. Vilan Vilan. Support vector machines used to estimate the battery state of charge. *IEEE Transactions on Power Electronics*, 28(12), December 2013.
- [19] B. Youn C. Hu and J. Chung. A multiscale framework with extended kalman filter for lithium-ion battery soc and capacity estimation. *Applied Energy*, 92:694–704, 2012.
- [20] S. Poll B. Saha, K. Goebel and J. Christophersen. Prognostics methods for battery health monitoring using a bayesian framework. *IEEE Transactions on Instrumentation and Measurement*, 58(2):291–296, 2009.
- [21] B. Saha K. Goebel M. Orchard, L. Tang and G. Vachtsevanos. Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices. *Studies in Informatics and Control*, 19:209–218, September 2010.
- [22] M. Cerda. Estimacion en linea del tiempo de descarga de baterias de ion-litio utilizando caracterizacion del perfil de utilizacion y metodos secuenciales de monte carlo. M.s thesis, Universidad de Chile, Santiago, Chile, 2012.
- [23] I. Roychoudhury S. Saha B. Saha A. Saxena, J. R. Celaya and K. Goebel. Designing data-driven battery prognostic approaches for variable loading profiles: Some lessons learned. In *European Conference of Prognostics and Health Management Society 2012*, 2012.
- [24] B. E. Olivares M. E. Orchard, M. A. Cerda and J. F. Silva. Sequential monte carlo methods for discharge time prognosis. *International Journal of Prognostics and Health Management*,

010:2153–2648, 2012.

- [25] S.H. Chung K.T. Chui, K.F. Tsang and L.F. Yeung. Appliance signature identification solution using k-means clustering. In *Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE*, pages 8420 – 8425, November 2013.
- [26] R. Xu and D. C. Wunsch II. *Clustering*, chapter Partitional Clustering, pages 63 – 69. John Wiley & Sons, Inc., Hoboken, New Jersey, 2009.
- [27] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*. Pearson Education, Inc, Boston, 4 edition, 2010.
- [28] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, chapter Maximum Likelihood Estimation, pages 157–173. Prentice Hall, Inc, New Jersey, 2010.
- [29] Todd K. Moon and Wynn C. Stirling. *Mathematical Methods and Algorithms*, chapter Estimation Theory, pages 542–590. Prentice Hall, Inc, 2000.
- [30] S. L. Myers R. E. Walpole, R. H. Myers and K. Ye. *Probability & Statistics for Enginners & Scientists*. Pearson Education, Inc, Boston, 2012.
- [31] A. M. Mood. *Intoduction to the Theory of Statistics*, chapter Test of Hyphoteses, pages 401–418. McGraw Hill, Inc., 1974.
- [32] G. Lugosi S. Boucheron and O. Bousquet. Concentration inequalities. *Advanced Lectures on Machine Learning, Lecture Notes in Computer Science*, 3176:208–240, 2004.
- [33] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, March 1963.
- [34] S. M. Ross. *Stochastic Processes*, chapter Markov Chains, pages 163 – 230. John Wiley & Sons, Inc, Boston, 2nd edition, 1996.
- [35] J. Norris. *Markov Chains*, chapter Discrete-time Markov chains, pages 1–59. Cambridge University Press, Nueva York.
- [36] M. Muhlich. Particle filters an overview, 2003.
- [37] S. Godsill A. Doucet and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10:197–208, 2000.
- [38] N. Gordon M. Sanjeev Arulampalam, S. Maskell and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.
- [39] N. Bergman. *Recursive Bayesian Estimation, Navigation and Tracking Aplications*. PhD thesis, Sweden, 1999.
- [40] T. A. a. L. Goodman. Statistical inference about markov chains. *Ann. Math. Statist*, 28(1):89–

110, 1957.

- [41] T. W. Anderson. Probability models for analyzing time changes in attitudes. *Mathematical Thinking in the Social Sciences*, edited by Paul F. Lazarsfeld, The Free Press, Glencoe, Illinois, 1954.
- [42] J. S. Hunter. The exponentially weighted moving average. *J. Qual. Technol.*, 18:203–209, 1986.
- [43] J. Liu and M. West. *Sequential Monte Carlo Methods in Practice*, chapter Combined Parameter and State Estimation in Simulation-Based Filtering, pages 197–223. Springer, 2001.
- [44] F. Tobar M. Orchard and G. Vachtsevanos. Outer feedback correction loops in particle filtering-based prognostic algorithms: Statistical performance comparison. *Studies in Informatics and Control*, 18:295–304, December 2009.
- [45] M. Chen and G. Rincon-Mora. Accurate electrical battery model capable of predicting runtime and i-v performance. *IEEE Transactions on Energy Conversion*, 21(2):504–511, 2006.
- [46] K. Bongort S. Engel, B. Gilmartin and A. Hess. Prognostics, the real issues involved with predicting life remaining. In *Aerospace Conference Proceedings, 2000 IEEE*, volume 6, pages 457–469, 2000.
- [47] G. Vachtsevanos C. Chen and M. Orchard. Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering. *IEEE Transactions on Industrial Electronics*, 58(9):4353–4364, September 2011.
- [48] M. Orchard B. Olivares, M. Cerda and J. Silva. Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. *IEEE Transactions on Instrumentation and Measurement*, 62(2), February 2013.