



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

OPERACIÓN ÓPTIMA DE SISTEMAS DE ALMACENAMIENTO BASADA EN
SDDP CON REPRESENTACIÓN MARKOVIANA DE LA INCERTIDUMBRE DE
PRECIOS DE LA ENERGÍA

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO

RODRIGO ANDRÉS VIDAL CARRASCO

PROFESOR GUÍA:
RODRIGO MORENO VIEYRA

MIEMBROS DE LA COMISIÓN:
ANGELA FLORES QUIROZ
MATÍAS NEGRETE PINCETIC

Este trabajo ha sido parcialmente financiado por: ANID-Subdirección de Capital
Humano/Magíster Nacional/2022 - 22221612, ANID Chile -
ANID/FONDECYT/Regular 1181928, ANID Chile - ANID/FONDECYT 1190228, ISCI
- ANID PIA/Apoyo AFB180003

SANTIAGO DE CHILE
2024

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN ELÉCTRICA Y MEMORIA PARA OPTAR AL
TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: RODRIGO ANDRÉS VIDAL CARRASCO
FECHA: 2024
PROF. GUÍA: RODRIGO ANDRÉS MORENO VIEYRA

OPERACIÓN ÓPTIMA DE SISTEMAS DE ALMACENAMIENTO BASADA EN
SDDP CON REPRESENTACIÓN MARKOVIANA DE LA INCERTIDUMBRE DE
PRECIOS DE LA ENERGÍA

Con la expansión de las energías ERNC y la transición hacia matrices más limpias y sostenibles, los desafíos para mantener una operación segura y confiable aumentan. Los sistemas de almacenamiento se posicionan como una de las opciones más completas para incrementar la flexibilidad de los SEP, pero para ello es fundamental contar con metodologías óptimas para la operación de estos dispositivos. Este trabajo estudia la aplicación de modelos de optimización multi-etapa con representación markoviana de la incertidumbre para operar sistemas BESS, utilizando MC-SDDP. Específicamente, se estudia el problema de arbitraje de energía, con y sin participación en mercados de reservas, bajo incertidumbre en los precios de la energía. Mediante comparación con otras metodologías de toma de decisión, se evalúa la política de operación derivadas del modelamiento propuesto. Los resultados muestran que la representación markoviana de la incertidumbre captura mejor la dinámica de precios de energía que una representación sin correlaciones temporales (independiente). En arbitraje de energía puro, el modelamiento multi-etapa markoviano supera en utilidades al modelamiento multi-etapa independiente en al menos un 15% y al aprendizaje reforzado en un 11%. Sin embargo, el modelo multi-etapa markoviano presenta complejidades adicionales para implementarse, como altos tiempos de cálculo y dificultades de convergencia.

A mi familia.

Agradecimientos

En primer lugar, quiero expresar mi más profundo agradecimiento a mi familia, quienes me han brindado un apoyo incondicional a lo largo de toda mi vida. A mis padres, les agradezco su cariño, apoyo y dedicación. En especial, a mi madre, por enseñarme que con perseverancia todo se logra.

También, quiero agradecer a los amigos que la universidad me ha dado. A Francys, Javier y Pato, por llenar de risas y momentos inolvidables esas largas tardes de estudio. Con mucho cariño a Pía, Migue y Gabo, por contagiar de alegría cada segundo en la Salita. Su amistad es uno de los tesoros más valiosos que me llevo de esta etapa. Espero que nuestra amistad perdure por muchos años más.

Finalmente, quiero agradecer a mi comisión por su inestimable ayuda en el desarrollo de este trabajo. A Rodrigo Moreno, mi profesor guía, por su confianza y orientación; cada conversación y comentario fueron esenciales para el desarrollo de esta investigación. A Angela Flores y Matías Negrete, les agradezco su disposición y apoyo constante.

Este trabajo ha sido financiado por los fondos: ANID-Subdirección de Capital Humano/Magíster Nacional/2022 - 22221612, ANID Chile - ANID/FONDECYT/Regular 1181928, ANID Chile - ANID/FONDECYT 1190228 e Instituto Sistemas Complejos de Ingeniería - ANID PIA/Apoyo AFB180003.

Tabla de Contenido

Nomenclatura	xi
1. Introducción	1
1.1. Motivación	1
1.2. Hipótesis de investigación	2
1.3. Objetivos	2
1.3.1. Objetivo general	2
1.3.2. Objetivos específicos	2
1.4. Estructura del documento	3
2. Revisión de la literatura	4
2.1. Estado del arte	4
2.2. Contribuciones	8
3. Metodología	10
3.1. Descripción general	10
3.2. Supuestos del modelamiento	11
3.3. Modelo matemático	12
3.3.1. Función de Bellman	12
3.3.2. Restricciones de factibilidad de reservas	13
3.3.3. Restricción de operación	14
3.4. Estimación estadística de la incertidumbre de precios	15

3.4.1.	Desestacionalización de datos	15
3.4.2.	Representación estadística de la incertidumbre	16
3.5.	Algoritmo de resolución	16
3.5.1.	Formulación del problema multi-etapa	17
3.5.2.	Formulación del algoritmo MC-SDDP	17
3.5.3.	Simplificación en la formulación del algoritmo SDDP frente a inde- pendencias temporales	19
3.5.4.	Regularización	20
3.5.5.	Estrategia de toma de decisión bajo escenarios reales	20
3.6.	Entrenamiento y simulación con datos históricos	21
3.7.	Modelos de referencia	22
3.7.1.	Modelo multi-etapa independiente	22
3.7.2.	Modelo determinista	22
3.7.3.	Aprendizaje reforzado	23
4.	Resultados y análisis	25
4.1.	Descripción de los casos de estudio	25
4.2.	Datos de entrada	26
4.2.1.	Datos históricos de precio	26
4.2.2.	Parámetros del BESS	27
4.3.	Resultados	27
4.3.1.	Comportamiento temporal de los precios de la energía	27
4.3.2.	Arbitraje de energía mediante optimización (caso base)	28
4.3.3.	Arbitraje de energía mediante optimización (sin desestacionalización)	30
4.3.4.	Arbitraje de energía y participación en el mercado de reservas me- diante optimización	31
4.3.5.	Arbitraje de energía mediante aprendizaje reforzado	35
5.	Conclusión	40

Bibliografía	47
ANEXOS	48
Anexo A. Mercado eléctrico australiano	48
Anexo B. Resultados	50
B.1. Arbitraje de energía mediante optimización	50
B.2. Arbitraje de energía y participación en el mercado de reservas mediante optimización	52
B.3. Arbitraje de energía mediante aprendizaje reforzado	54

Índice de Tablas

3.1. Hiperparámetros de PPO.	23
3.2. Características utilizadas e hiperparámetros de la arquitectura.	23
4.1. Parámetros técnicos del BESS.	27
4.2. Gap promedio de los modelos de optimización multi-etapa utilizados en esta investigación.	33
4.3. Características e hiperparámetros óptimos.	35
A.1. Mercados australianos de servicios auxiliares de control de frecuencia. . . .	49
B.1. Tiempo de computo y uso de memoria de los modelos de toma de decisión tratados en esta investigación. Los valores indicados representan los recursos empleados durante una única ejecución de entrenamiento, y no deben interpretarse como valores promedio.	55

Índice de Ilustraciones

3.1. Diagrama del entrenamiento y prueba del modelo multi-etapa.	21
3.2. Diagrama del entrenamiento y prueba del modelo determinista.	22
3.3. Arquitectura de la red del actor y de la red evaluadora.	24
4.1. Curvas diarias del precio de la energía correspondientes al bus QLD y al año 2022. Los datos son submuestreados a una resolución de 30 minutos.	28
4.2. Modulo de la matriz de correlación por bus y año. Los datos son submuestreados a una resolución de 30 minutos.	28
4.3. Utilidades relativas de los modelos de optimización al arbitrar energía y desestacionalizar precios, mercado con resolución de 5 minutos. El error de estimación de la media considera un 95 % de confianza.	29
4.4. Utilidades relativas de los modelos de optimización al arbitrar energía. Se desestacionalizan los precios y se consideran diferentes resoluciones en la operación del mercado.	30
4.5. Utilidades relativas de los modelos basados en optimización al arbitrar energía, mercado con resolución de 5 minutos. Las barras representan las utilidades al no desestacionalizar precios. Los puntos representan las utilidades al desestacionalizar los precios de la energía y luego corregir las utilidades.	31
4.6. Utilidades relativas de los modelos basados en optimización al participar en el mercado de reservas y arbitrar de energía, mercado con resolución de 10 minutos y precios desestacionalizados. Se excluyen los días con utilidades superiores a 5.000[AUD].	32

4.7. Utilidades de los modelos de optimización al participar en el mercado de reservas y arbitrar de energía, mercado con resolución de 10 minutos y precios desestacionalizados. Se excluyen los días con utilidades superiores a 5.000[AUD]. Referencialmente, se presentan las utilidades de un modelo que revela la incertidumbre en la primera etapa (primera barra de cada estado).	34
4.8. Utilidades relativas de los modelos de optimización y el basado en aprendizaje reforzado al realizar arbitraje de energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.	36
4.9. Operación de los primeros 15 días del año 2022 al utilizar modelos basados en optimización para arbitrar energía. Datos del nodo de referencia ubicado en Queensland.	38
4.10. Operación de los primeros 15 días del año 2022 al utilizar aprendizaje reforzado para arbitrar energía. Datos del nodo de referencia ubicado en Queensland.	38
4.11. Distribución de las utilidades de los modelos basados en optimización y en aprendizaje reforzado al arbitrar energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.	39
B.1. Utilidades relativas de los modelos de optimización al arbitrar energía. No se desestacionalizan los precios y se consideran diferentes resoluciones en la operación del mercado.	51
B.2. Utilidades de los modelos basados en optimización al arbitrar energía, mercado con resolución de 5 minutos. Las barras representan las utilidades corregidas al usar precios desestacionalizados. Los puntos representan las utilidades al no desestacionalizar precios.	52
B.3. Precio promedio y rango interdecil horario de los precios de las reservas del año 2022. Se han excluido aquellas curvas de precio que presentan valores superiores 600[AUD/MW] o inferiores a -600 [AUD/MW].	53
B.4. Utilidades de los modelos de optimización al participar en el mercado de reservas y arbitrar energía, mercado con resolución de 10 minutos y precios desestacionalizados. Se excluyen los días con utilidades superiores a 10.000[AUD].	53
B.5. Utilidades de los modelos de optimización y el basado en aprendizaje reforzado al realizar arbitraje de energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.	54

B.6. Operación de los primeros 15 días del año 2022 al utilizar modelos basados en optimización para arbitrar energía. Datos del nodo de referencia ubicado en Tasmania.	56
B.7. Operación de los primeros 15 días del año 2022 al utilizar aprendizaje reforzado para arbitrar energía. Datos del nodo de referencia ubicado en Tasmania.	56

Nomenclatura

Abreviaturas

BESS	Sistema de almacenamiento de energía por baterías (por sus siglas en inglés).
ERNC	Energías renovables no convencionales.
ESS	Sistema de almacenamiento de energía (por sus siglas en inglés).
FCAS	Servicio auxiliar de control de frecuencia (por sus siglas en inglés).
MDP	Proceso de Decisión de Markov (por sus siglas en inglés).
NEM	Mercado eléctrico nacional australiano (por sus siglas en inglés).
NOB	Banda de operación normal (por sus siglas en inglés). Desde 49,85 a 50,15Hz.
NSW	Abreviatura del nodo de referencia ubicado en New South Wales.
OPF	Flujo óptimo de potencia (por sus siglas en inglés).
PPO	Optimización de Política Proximal (por sus siglas en inglés).
QLD	Abreviatura del nodo de referencia ubicado en Queensland.
SA	Abreviatura del nodo de referencia ubicado en South Australia.
SDDP	Programación dinámica dual estocástica (por sus siglas en inglés).
SEP	Sistema eléctrico de potencia.
SOC	Estado de carga (por sus siglas en inglés).
TAS	Abreviatura del nodo de referencia ubicado en Tasmania.
VIC	Abreviatura del nodo de referencia ubicado en Victoria.
VPP	Planta de potencia virtual (por sus siglas en inglés).

Índices y conjuntos

t	Índice de las etapas o bloques de operación.
m	Índice de los mercados de reservas.

Parámetros

M	Número de mercados de reserva ante contingencias.	-
\bar{P}^c	Potencia máxima de carga.	MW
\bar{P}^d	Potencia máxima de descarga.	MW
\bar{S}	Nivel de carga máximo admisible.	MWh
\underline{S}	Nivel de carga mínimo admisible.	MWh
T	Número de bloques de operación.	-
ΔT	Resolución temporal del modelo de optimización.	h
ΔT_m	Tiempo durante el que se debe sostener el suministro de las reservas del mercado m .	h

η^c	Eficiencia energética de carga.	-
η^d	Eficiencia energética de descarga.	-
π_t	Precio de la energía en el bloque t .	AUD/MWh
$\pi_{t,m}^L$	Precio de la reserva de bajada del mercado m en el bloque t .	AUD/MW
$\pi_{t,m}^R$	Precio de la reserva de subida del mercado m en el bloque t .	AUD/MW

Variables

P_t^c	Potencia de carga en el bloque t .	MW
$P_j^{c,\text{aux}}$	Potencia de carga factible en j bloques a partir del actual.	MW
P_t^d	Potencia de descarga en el bloque t .	MW
$P_j^{d,\text{aux}}$	Potencia de descarga factible en j bloques a partir del actual.	MW
$q_{t,m}^L$	Reserva de bajada comprometida en el mercado m y bloque t .	MW
$q_{t,m}^R$	Reserva de subida comprometida en el mercado m y bloque t .	MW
\mathbf{q}	Vector que contiene las reservas comprometidas para el horizonte de operación, $\mathbf{q} \in \mathbb{R}^{T \times 2M}$.	MW
s_t	Nivel de carga al final del bloque t .	MWh
s_j^{aux}	Nivel de carga factible en j bloques a partir del actual.	MWh

Capítulo 1

Introducción

1.1. Motivación

Históricamente, el suministro eléctrico mundial ha dependido en gran medida de combustibles fósiles como el carbón, diésel y gas. Sin embargo, debido a la contaminación generada por estas tecnologías y a la urgencia de abordar la crisis climática global, se han impulsado iniciativas en todo el mundo para reemplazar gradualmente la generación convencional por fuentes de energía más limpias y sostenibles. Así, en los últimos años, la generación a través de energías renovables no convencionales (ERNC) ha experimentado un crecimiento sin precedentes, concentrando inversiones en generación fotovoltaica y eólica [1–3].

No obstante, debido a la alta variabilidad e intermitencia de las plantas de generación renovables, la transición de una matriz energética basada principalmente en generación convencional a una basada en ERNC plantea una serie de desafíos en términos de seguridad y confiabilidad [4–6]. En consecuencia, se vuelve imperativo flexibilizar los sistemas eléctricos, instalando equipos e implementando estrategias que permitan tener una operación adaptable a la variabilidad e incertidumbre de la generación y la demanda.

En este contexto, los sistemas de almacenamiento de energía (ESS, por sus siglas en inglés), y en particular, los sistemas de almacenamiento de energía por baterías (BESS, por sus siglas en inglés), han surgido como una de las soluciones más versátiles para otorgar a los sistemas eléctricos la flexibilidad requerida. Su rápida respuesta, junto con sus capacidades para proveer servicios de reserva, regulación de frecuencia, arbitraje de energía, gestión de la demanda y/o generación, pueden contribuir significativamente a una operación con altos porcentajes de generación renovable [7–11]. Sin embargo, para realizar inversiones en este tipo de equipos, es fundamental contar con metodologías de operación que posibiliten un uso eficiente de los mismos.

Entre las herramientas utilizadas para la operación óptima de sistemas de almacenamiento, los modelos de optimización estocásticos, y en particular los modelos de optimización estocásticos multi-etapa, son los más comunes [12–17]. Estos modelos gestionan

la incertidumbre inherente a la generación renovable y a los mercados de energía en situaciones donde esta se revela de manera secuencial. Sin embargo, a menudo suponen independencia temporal en la representación de la incertidumbre para facilitar la implementación de metodologías de resolución y descomposición eficientes [18,19]. Este supuesto de independencia en la incertidumbre no siempre se cumple, ya que las variables aleatorias involucradas en la operación de sistemas eléctricos de potencia, como los niveles de generación renovable y otras variables de los mercados, a menudo presentan correlaciones temporales [20,21].

Dentro del marco de los modelos de optimización estocásticos multi-etapa para la operación de sistemas de almacenamiento, esta investigación explora los beneficios económicos obtenidos al incorporar correlaciones temporales en el modelamiento de la incertidumbre. En particular, se aborda específicamente el problema de arbitraje de energía considerando que la incertidumbre proviene de los precios de la energía eléctrica.

1.2. Hipótesis de investigación

La hipótesis de investigación propuesta en este trabajo sostiene que, en el contexto de la operación de sistemas BESS y cuando la incertidumbre radica en las fluctuaciones de los precios de la energía, los modelos de optimización multi-etapa que adoptan una representación markoviana de la incertidumbre generan políticas de operación más eficientes que los modelos multi-etapa con representación independiente de la incertidumbre y que los modelos no dinámicos (estáticos) usados en la actualidad.

1.3. Objetivos

A continuación se enuncian los objetivos generales y específicos que se buscan alcanzar con el desarrollo de esta tesis.

1.3.1. Objetivo general

Proponer un modelo de optimización multi-etapa con representación markoviana de la incertidumbre, en el contexto de la operación de un BESS, y evaluar la eficacia en la toma de decisiones de este, contrastándola con otras metodologías basadas tanto en optimización como en aprendizaje reforzado.

1.3.2. Objetivos específicos

- Formular modelos matemáticos de optimización multi-etapa para la operación de un BESS, que utilicen una representación markoviana e independiente de la incer-

tidumbre, destinados al arbitraje de energía y la participación en el mercado de reservas.

- Implementar una estrategia para resolver eficientemente los problemas de optimización multi-etapa utilizando programación dinámica dual estocástica (SDDP, por sus siglas en inglés).
- Implementar una estrategia para la operación de un BESS utilizando aprendizaje reforzado.
- Analizar comparativamente las políticas de operación derivadas de los mecanismos de toma de decisión mediante simulación con datos reales.

1.4. Estructura del documento

En el segundo capítulo de este documento incluye una revisión bibliográfica que detalla el estado del arte en operación y planificación de equipos BESS. Adicionalmente, este capítulo expone las principales contribuciones de la investigación. En el tercer capítulo se introduce la metodología, cuyo propósito es formular el modelo matemático de optimización y describir el procedimiento empleado para simular la operación con dicho modelo. En el cuarto capítulo se presentan las características generales de cada caso de estudio, los datos de entrada utilizados en las pruebas y los resultados obtenidos al simular escenarios de operación. Finalmente, en el quinto capítulo se presentan las principales conclusiones obtenidas del trabajo y se proponen posibles líneas de investigación para el futuro.

Capítulo 2

Revisión de la literatura

2.1. Estado del arte

A medida que diversos países alrededor del mundo avanzan en la adopción de energías renovables no convencionales y establecen objetivos de descarbonización [2], sus sistemas eléctricos enfrentan crecientes desafíos en la planificación y operación [22]. El aumento en las cuotas de generación renovable conlleva mayores dificultades en la predicción de generación, reducciones en los niveles de inercia y desafíos para mantener una operación estable tanto en voltaje como en ángulo [4]. Esto hace imperativo el desarrollo de estrategias para asegurar redes flexibles, confiables y resilientes [23]. En este escenario, los sistemas de almacenamiento se perfilan como la opción más prometedora para habilitar las redes eléctricas para una generación renovable en aumento [24–26]. Sin embargo, para una adopción masiva de estos dispositivos, es esencial desarrollar modelos que optimicen los beneficios derivados de su operación.

En el contexto de la operación óptima de sistemas de almacenamiento, [27] y [28] presentan modelos de planificación que evalúan la inversión en sistemas BESS. Dichos estudios abordan la participación en el mercado de energía intradiario y la provisión de servicios de control de frecuencia, sin tener en cuenta estocasticidades. De manera similar, [29] introduce un modelo de flujo óptimo de potencia (OPF, por sus siglas en inglés) para coordinar la operación de prosumidores que poseen sistemas de almacenamiento y paneles solares. Este último tampoco incorpora un enfoque estocástico en ninguna de las variables del problema.

Diferente de los estudios que ya mencionados, la literatura también ha evidenciado un crecimiento en la propuesta de modelos de optimización que incluyen enfoques estocásticos. Estos modelos, en su esencia, tratan de manejar la incertidumbre inherente a la generación renovable y los mercados de energía. Podemos clasificarlos ampliamente en dos grandes categorías: modelos de dos etapas y modelos multi-etapa.

Los modelos de dos etapas son idóneos para situaciones en las que una primera decisión se toma bajo incertidumbre. Una vez que esta incertidumbre se revela, se toma una

segunda decisión con información completa. Dentro de esta categoría, los estudios principalmente proponen modelos de planificación, orientados a determinar el nivel óptimo de inversión en sistemas de almacenamientos mientras se considera la incertidumbre de variables como la generación [30–32] o la generación y demanda [33,34]. En estos trabajos, se toma una decisión de inversión en infraestructura ante incertidumbre operativa, y tras esa decisión, el sistema se opera con conocimiento de las condiciones reales.

A pesar que los modelos de dos etapas, per se, no permiten políticas dinámicas adaptativas de toma de decisiones, ciertos estudios han explorado su aplicación en contextos de operación. Por ejemplo, en [12] se formula un modelo de dos etapas que elabora un esquema de cambio de servicio provisto por un BESS (arbitraje de energía ó regulación de frecuencia), considerando incertidumbre en los precios de la energía y desgaste del equipo. Por otro lado, [13] aborda la operación de un BESS en una microrred de corriente continua, enfrentando incertidumbre en la generación y demanda, mediante un modelo entero-mixto no lineal. En este enfoque, la primera etapa se centra en la operación desde una perspectiva energética, y, una vez revelada la incertidumbre en la segunda etapa, se tratan los flujos en la red buscando minimizar las pérdidas.

En contraste, los modelos multi-etapa se diseñan para situaciones de incertidumbre multi-etapa donde la toma de decisiones y el revelado de la incertidumbre ocurre secuencialmente. En [35] se propone un OPF multi-etapa para operar un SEP con instalación de dispositivos de almacenamiento y generación renovable de naturaleza estocástica. Otro estudio relevante es [36], que aborda la planificación y operación de mantenimientos de plantas hidroeléctricas, incorporando estocasticidad tanto de los precios de la energía como en variables hidrológicas. Finalmente, existe una serie de investigaciones que se centran en la optimización de las ofertas para mercados day-ahead. Estos modelos permiten reajustes posteriores en la programación a través de mercados intradiarios y suelen adoptar la perspectiva de un operador de planta virtual (VPP) [21,37].

Dada la complejidad computacional inherente a los modelos multi-etapa, es fundamental contar con algoritmos de resolución especializados para resolverlos. Estos algoritmos deben ser capaces de manejar una vasta cantidad de escenarios, garantizando decisiones óptimas. En este sentido, a la hora de abordar problemas lineales multi-etapa, los algoritmos basados en SDDP [18,19] son los más recurrentes en la literatura. Estos algoritmos aprovechan la estructura lineal del problema y la independencia de la incertidumbre entre etapas para descomponerlo y tratarlo eficientemente mediante cortes de Benders. Dado que la incertidumbre en cada etapa no depende de la realización de la incertidumbre en la etapa previa, el problema de decisión en cada etapa se reduce a un único problema de optimización. Para construir los cortes, se simula la evolución de la variable estocástica y se resuelven los problemas de optimización de cada etapa, desde la etapa inicial hasta la final, recopilando la información necesaria para construir los cortes y poblar las restricciones del problema de cada etapa.

Originalmente, el algoritmo SDDP se diseñó para resolver problemas multi-etapa estocásticos lineales sin aversión al riesgo. Sin embargo, con el tiempo, se han introducido adaptaciones para manejar problemas más desafiantes. Por ejemplo, en [38,39] se presenta una variante del algoritmo que incorpora aversión al riesgo. En cambio, en [40,41]

se propone MC-SDDP, una adaptación con un enfoque markoviano de la incertidumbre. En esta variante, la distribución de las futuras variables aleatorias está condicionada a las realizaciones previas, marcando una diferencia con las metodologías anteriores que no tomaban en cuenta correlaciones temporales.

Diversas investigaciones se han enfocado en acelerar la rapidez con la que el algoritmo SDDP converge. Por ejemplo, en [42, 43] se examinan técnicas de regularización que potencian la convergencia del algoritmo. Por otro lado, [44] propone una adaptación que permite una ejecución en paralelo y [45] propone una variante para facilitar la ejecución distribuida. Otras áreas de investigación se han centrado en adaptar el algoritmo SDDP para resolver problemas multi-etapa entero-mixtos, estableciendo garantías de convergencia tras una cantidad finita de iteraciones [46, 47]. De manera similar, [48] modifica el algoritmo para garantizar la convergencia en ciertos problemas multi-etapa estocásticos entero-mixtos no lineales. En [49], se presenta una notable integración entre SDDP y aprendizaje reforzado, utilizando redes neuronales para estimar la función de costos futuros, delineando así una versión de aprendizaje reforzado basada en modelo.

El uso principal de los algoritmos SDDP ha sido en coordinación hidro-térmica. Esta área se enfoca en minimizar los costos operativos de sistemas eléctricos que combinan generadoras hidroeléctricas con embalses, que enfrentan afluentes estocásticos, y plantas térmicas [14–16, 50]. Hay investigaciones que amplían el alcance de esta coordinación para considerar aspectos como requerimientos de riego, restricciones de navegación, pesca y caudales ecológicos [51, 52]. La capacidad del algoritmo para resolver problemas multi-etapa de administración de recursos también lo hace idóneo para resolver otros problemas relacionados con almacenamiento.

En [17], se presenta un modelo para operar un sistema de almacenamiento de energía que tiene en cuenta el desgaste del equipo y modela la incertidumbre en la generación eólica y solar, así como en la demanda. Por otro lado, en [53], se expone un modelo destinado a la gestión energética en una comunidad residencial que utiliza sistemas de almacenamiento y generación fotovoltaica. Este modelo incorpora la incertidumbre tanto en la generación como en la demanda, además de incluir el dimensionamiento de la inversión en almacenamiento. En [54], se propone un modelo para la gestión de instalaciones de almacenamiento de gas natural. En este caso, el trabajo tiene en cuenta la incertidumbre en el precio del gas en el mercado spot y en el futuro. El objetivo principal es determinar el régimen óptimo de compra y venta, considerando las restricciones en el almacenamiento. Además, en otras investigaciones que se alejan del área de energía, se ha utilizado el algoritmo en el ámbito del retail [55], desarrollando estrategias de gestión de mercancías bajo incertidumbre de precios.

En el ámbito de los modelos que adoptan una representación markoviana de la incertidumbre, [56] introduce una modificación al algoritmo SDDP. Esta adaptación permite modelar transiciones entre las distribuciones asociadas a diferentes hidrologías, tales como períodos secos, normales y húmedos, en el contexto del problema de coordinación hidro-térmica. De manera similar, [40, 41] utilizan MC-SDDP y evalúan su eficacia en pruebas relacionadas con el despacho hidro-térmico y la incertidumbre en los afluentes. En particular, en [41], los autores sugieren como trabajo futuro explorar las capacidades

del modelamiento MC-SDDP para incorporar incertidumbre en precios (i.e., aleatoriedad en los coeficientes de la función objetivo). A raíz de esto, en [54] se propone el modelo de gestión de almacenamiento de gas con incertidumbre en los precios del gas en el mercado spot comentado previamente. Por otro lado, [42, 43] exploran la convergencia y eficacia de las políticas operativas generadas mediante MC-SDDP en el despacho de un SEP que integra sistemas de almacenamiento de energía. En estas investigaciones, el recurso eólico es la principal fuente de incertidumbre y se analiza la convergencia del algoritmo al variar la cantidad de sistemas BESS en la infraestructura eléctrica. En [57], se emplea MC-SDDP para resolver el problema de unit commitment de un SEP con centrales de bombeo, considerando incertidumbre en la generación eólica. Para acelerar la resolución del problema, la implementación del algoritmo SDDP incorpora una variante de la regularización propuesta en [43]. Las políticas de operación obtenidas a partir del modelado propuesto se comparan con las de otras estrategias actualmente utilizadas en la industria. A la fecha de esta investigación, no se identifican estudios que incluyan MC-SDDP y la operación de sistemas de almacenamiento de energía utilizando incertidumbre en los precios de la energía, ya sea para arbitraje o para la provisión de servicios complementarios.

En los últimos años, como alternativa a los métodos tradicionales de operación basados en optimización, se han desarrollado investigaciones que exploran el uso del aprendizaje reforzado en el proceso de toma de decisiones. Estos modelos, aunque no necesariamente alcanzan óptimos globales, pueden identificar relaciones complejas en la operación de sistemas eléctricos y en los mercados a partir de datos históricos o simulaciones. En la recopilación presentada en esta tesis, se identificaron escasos trabajos que involucraran la utilización de aprendizaje reforzado en el contexto de la operación de sistemas eléctricos y ningún trabajo que compare esquemas de operación basados en optimización matemática con aquellos basados en aprendizaje reforzado, en el contexto de operación de sistemas de almacenamiento.

En [20] se emplea aprendizaje reforzado para operar en el mercado intradiario de Alemania desde la perspectiva del operador de un parque eólico, considerando solo la participación en el mercado de energía. Por otro lado, en [58] se propone un método para operar un BESS, realizando arbitraje de energía y participando en el mercado de regulación de frecuencia. Ambos estudios emplean la técnica de optimización de política proximal (PPO, por sus siglas en inglés) [59], destacada por permitir decisiones en espacios continuos y por su alta estabilidad durante el entrenamiento. En contraste, [60] estudia el uso de Q-learning [61] para arbitrar energía con un BESS. Esta técnica, una de las más populares en aprendizaje reforzado, opera en un espacio de decisiones discreto, implicando una representación segmentada de las posibles acciones a tomar.

En la revisión de la literatura, no se han identificado estudios que integren MC-SDDP con la operación de sistemas de almacenamiento de energía considerando incertidumbre en los precios de la energía, ya sea para arbitraje o para la provisión de servicios complementarios. Suponiendo que el sistema de almacenamiento es lo suficientemente pequeño para no influir en las tendencias de los precios, esta aproximación permitiría reducir la complejidad del problema de operación. Esto se lograría al centrarse en el seguimiento de una única señal, la de precios, en lugar de modelar completamente el resto del sistema eléctrico y las variables estocásticas que influyen en su operación, como mantenimientos,

fallas, fluctuaciones en la demanda y volatilidades en los precios de los combustibles, entre otras.

En la literatura tampoco se han comparado las metodologías convencionales de operación de sistemas de almacenamiento basadas en optimización matemática con estrategias emergentes como el aprendizaje reforzado. Mucho menos se ha realizado una comparación entre el MC-SDDP, una metodología prometedora de optimización matemática que ofrece una representación detallada de la incertidumbre, y el aprendizaje reforzado.

2.2. Contribuciones

Con la creciente expansión de las energías renovables, la participación en los mercados intradiarios se ha convertido en un mecanismo esencial para sobrellevar la volatilidad introducida en la generación de energía. En este sentido, los BESS emergen como una oportunidad debido a que su rápida capacidad de respuesta posibilita ajustes dinámicos en tiempo real. Sin embargo, la mayoría de las investigaciones abordan la operación de estos equipos sin considerar estocasticidad o considerando incertidumbre de dos etapas. Por esta razón, los modelos propuestos a menudo no son aplicables directamente en contextos prácticos de operación multi-etapa con incertidumbre operacional.

En este trabajo se aborda el problema del arbitraje de energía de un BESS utilizando MC-SDDP. La novedad de esta propuesta radica en que la fuente de incertidumbre proviene de los precios de la energía, a diferencia de los enfoques tradicionales que se centran en la disponibilidad del recurso renovable. Aunque esta aproximación requiere un mayor esfuerzo para representar la variable estocástica, ya que es el resultado de múltiples procesos que afectan la operación de un SEP (como mantenimientos, fallas, fluctuaciones en la demanda, volatilidades en los precios de los combustibles, variabilidad de la generación y variables meteorológicas), y necesita herramientas de regularización para eliminar posibles estacionalidades en datos históricos, permite reducir la complejidad del problema de operación. Esto se debe a que, al centrar la operación en la señal de precios, que incide directamente en las utilidades percibidas, se pueden omitir otras variabilidades que afectan a los sistemas eléctricos. De esta manera, se facilita la obtención de políticas óptimas de operación mediante la simplificación del modelo de optimización, adoptando un enfoque tomador de precios.

Para evaluar la eficacia de la representación de la incertidumbre, se resuelve el mismo problema bajo dos escenarios adicionales alternativos: uno sin modelamiento estocástico y otro que asume una representación de la incertidumbre sin correlaciones temporales. Este análisis permite juzgar la política de operación derivada de cada representación de la incertidumbre, aspecto usualmente pasado por alto en las investigaciones.

Adicionalmente, también se utiliza MC-SDDP para abordar el problema de arbitraje de energía y participación en el mercado de reservas. Este problema, a pesar de integrar la toma de decisiones de primera etapa (oferta de reservas) y la toma dinámica de decisiones (arbitraje en tiempo real), ha sido escasamente estudiado en la literatura. Es probable que

esto se deba el alto costo computacional de MC-SDDP y a los desafíos de convergencia de la metodología. Sin embargo, desde la perspectiva de toma de decisiones, es uno de los enfoques más precisos que se dispone actualmente.

Como parte final de la investigación, se aborda el problema de arbitraje de energía aplicando aprendizaje reforzado, específicamente en contextos sin participación en el mercado de reservas. El objetivo de este análisis es comparar el rendimiento de MC-SDDP respecto a los modelos actuales basados en aprendizaje reforzado.

Esta comparación es esencial. Los modelos de optimización, diseñados para encontrar soluciones óptimas globales, se basan en una representación matemática de los fenómenos involucrados (exacta o aproximada), lo cual es útil cuando existe una comprensión acabada del problema. No obstante, estos modelos a menudo requieren suposiciones y simplificaciones que pueden limitar su aplicabilidad en situaciones no convexas, no lineales, o complejas de representar (por ejemplo, incertidumbre en precios de la energía). Por otro lado, los modelos de aprendizaje reforzado son altamente flexibles y, usualmente, no necesitan suposiciones previas detalladas, adaptándose a relaciones complejas y no lineales mediante el aprendizaje de experiencias acumuladas. Aunque adecuados para situaciones difíciles de modelar explícitamente, son intensivos en datos y recursos computacionales, y no garantizan soluciones óptimas globales.

Comparar ambas metodologías permite identificar las fortalezas y debilidades de cada enfoque, facilitando la elección de la metodología más adecuada para la operación de sistemas BESS. Este aspecto, no explorado en la literatura, es crucial para una correcta toma de decisiones.

Capítulo 3

Metodología

En este capítulo se presenta la formulación matemática del modelo de optimización, el algoritmo de resolución escogido para resolverlo y otras metodologías de toma decisiones con las que se comparará su efectividad. Se comienza ofreciendo una descripción general del modelo, destacando sus principales características. A continuación, se listan los principales supuestos del modelamiento. Después, se presenta formalmente el modelo matemático, el procedimiento de estimación de parámetros, el algoritmo de resolución y la metodología de simulación. Finalmente, se exponen algunos modelos de optimización y de machine learning que se utilizarán para comparar y evaluar el rendimiento del modelo propuesto.

3.1. Descripción general

Se propone un modelo basado en optimización matemática destinado a la operación de un BESS stand-alone. Dicho modelo, de carácter estocástico y multi-etapa, tiene por objetivo maximizar las utilidades al participar en el mercado de reservas ante contingencias y realizar arbitraje de energía.

Como regla general, la inclusión de una batería stand-alone en un sistema eléctrico podría tener un impacto significativo en los precios de las reservas y de la energía. En consecuencia, una simulación precisa de la dinámica de precios en el mercado requiere un modelo detallado donde una gran variedad de unidades pueden presentar ofertas. Dado que este tipo de modelo tiene un alto costo computacional y supera los alcances de esta investigación, en este trabajo se asume que el BESS actúa como un tomador de precios, tanto en el mercado de reservas como en el de energía.

Con respecto al funcionamiento específico del modelo, este supone un conocimiento completo de los precios de las reservas, mientras que los precios de la energía constituyen la fuente de incertidumbre del modelamiento. De esta forma, durante la primera etapa se decide un monto de reservas que maximice conjuntamente los ingresos obtenidos al ofertar reservas y los ingresos esperados al operar en tiempo real. Para cada una de las etapas

subsiguientes, se maximizan las utilidades obtenidas al arbitrar energía en esa etapa y los ingresos esperados por futuros arbitrajes, siempre considerando la cantidad de reservas comprometidas en la etapa inicial.

El aspecto novedoso del modelamiento radica en que la incertidumbre de los precios de la energía se modela considerando correlaciones temporales. Esto se logra al modelar los precios mediante un proceso markoviano, donde el precio de la energía depende del precio en la etapa anterior.

Por último, un aspecto no considerado en el modelamiento propuesto es el pago por la utilización de reservas. El modelo se centra únicamente en los pagos por la disponibilidad de reservas y por arbitraje de energía. La justificación para esta decisión se fundamenta en que, en la práctica, la utilización de reservas tiene una baja probabilidad de ocurrencia, y por lo tanto, es despreciable su efecto en términos de operación energética e ingresos cuando el horizonte es breve [28].

3.2. Supuestos del modelamiento

A continuación, se presentan los supuestos del modelamiento matemático desarrollado:

- El BESS es suficientemente pequeño para no influir de manera significativa en la dinámica de precios del mercado, razón por la cual se le considera un tomador de precios.
- El modelo propuesto no tiene en cuenta un pago por parte del operador del BESS hacia el operador de mercado en caso de ser uno de los causantes de los requerimientos de reservas.
- La participación en el mercado de reservas considera un conocimiento completo de los precios de las reservas en el horizonte de evaluación.
- El modelo propuesto no contempla la ocurrencia de fallas, por tanto no existe un pago al suministrar reservas.
- El modelo no toma en cuenta el desgaste del BESS ni ninguna penalización asociada a la cantidad de ciclos de carga realizados durante la operación.

3.3. Modelo matemático

3.3.1. Función de Bellman

La ecuación (3.1) muestra la función de Bellman correspondiente a la primera etapa¹ del problema de optimización. Esta función maximiza los ingresos por compromiso de reservas y los ingresos esperados futuros. La función Q_1 representa los ingresos esperados resultantes para la decisión de reservas \mathbf{q} , el estado inicial de carga s_0 y el precio de la energía en la primera etapa de operación π_1 . Adicionalmente, $\mathcal{F}_0(\mathbf{q}, s_0)$ es el conjunto de restricciones que modela las reservas que es factible suministrar dada la carga inicial.

$$\begin{aligned}
 Q_0(s_0) = \max \quad & \sum_{\substack{\forall t \in (1, \dots, T) \\ \forall m \in (1, \dots, M)}} \left(\pi_{t,m}^L \cdot q_{t,m}^L + \pi_{t,m}^R \cdot q_{t,m}^R \right) \cdot \Delta T + \mathbb{E}_{\pi_1} [Q_1(\mathbf{q}, s_0, \pi_1)] \\
 \text{s.a.} \quad & \mathcal{F}_0(\mathbf{q}, s_0) \\
 & \mathbf{q} \geq 0
 \end{aligned} \tag{3.1}$$

En la ecuación (3.2) se presenta la función de Bellman de las etapas de operación. Esta función tiene por objetivo maximizar tanto los ingresos presentes, obtenidos al arbitrar energía, como los ingresos futuros esperados. A diferencia de la etapa inicial, los ingresos futuros esperados en estas etapas se calculan teniendo en cuenta la correlación en los precios de la energía. El conjunto de restricciones $\mathcal{F}_t(\mathbf{q}, s_t)$ determina los estados de carga que son consistentes con el suministro de reservas en el futuro. Por otro lado, el conjunto $\mathcal{O}_t(\mathbf{q}, s_{t-1}, P_t^d, P_t^c, s_t)$ modela las decisiones de operación (P_t^d , P_t^c y s_t) que son compatibles con la decisión de reservas \mathbf{q} y el estado de carga en la etapa anterior s_{t-1} . Dado que el modelo de optimización considera T etapas de operación, $Q_{T+1} = 0$ y $\mathcal{F}_T = \emptyset$.

$$\begin{aligned}
 Q_t(\mathbf{q}, s_{t-1}, \pi_t) = \max \quad & \pi_t \cdot (P_t^d - P_t^c) \cdot \Delta T + \mathbb{E}_{\pi_{t+1}|\pi_t} [Q_{t+1}(\mathbf{q}, s_t, \pi_{t+1})] \\
 \text{s.a.} \quad & \mathcal{F}_t(\mathbf{q}, s_t) \\
 & \mathcal{O}_t(\mathbf{q}, s_{t-1}, P_t^d, P_t^c, s_t) \\
 & P_t^d, P_t^c, s_t \geq 0
 \end{aligned} \tag{3.2}$$

¹Generalmente, en la literatura, el término “primera etapa” se refiere a la etapa inicial de un problema multi-etapa. No obstante, en esta investigación, para simplificar la notación, las etapas se indexan desde cero. Por consiguiente, la “primera etapa” es la etapa $t = 0$, y las etapas de operación están indexadas desde $t = 1$ hasta $t = T$.

3.3.2. Restricciones de factibilidad de reservas

Para garantizar que la decisión de reservas tomada en la primera etapa no comprometa la factibilidad de la operación en las etapas posteriores, se introducen restricciones que exigen la existencia de al menos una forma de operar compatible con la decisión de reservas y el estado inicial de carga. En caso de no introducir estas restricciones en el problema de la primera etapa, podría elegirse una combinación de reservas no alcanzable por las restricciones de operación del BESS. Por ejemplo, si la batería inicia descargada, la decisión de reservas no puede considerar la provisión de reservas de subida en el primer bloque de operación.

Asimismo, la decisión de carga y descarga en cada etapa posterior debe garantizar que el nuevo estado de carga sea factible para la provisión de reservas en las etapas siguientes. Estas restricciones, representadas por el conjunto $\mathcal{F}_t(\mathbf{q}, s_t)$, modelan el funcionamiento de la batería desde la etapa $t+1$ hasta el final del horizonte de operación, comenzando con un estado de carga s_t , como se muestra en la ecuación (3.3). De este modo, las restricciones garantizan, al menos, una secuencia de estados de carga $s_t^{aux}, \dots, s_T^{aux}$ compatibles con la decisión de reservas de la primera etapa \mathbf{q} , tal que $s_t = s_t^{aux}$.

Las restricciones (3.3a) y (3.3b) aseguran que la potencia de carga y descarga permanezca dentro del rango de operación nominal. Además, durante contingencias, el BESS tiene que ser capaz de aumentar o disminuir la inyección de potencia desde su estado pre-falla, suministrando las cantidad de reservas comprometida, restricción (3.3c) y (3.3d). Cabe destacar que en estas restricciones se asume la activación simultánea de las reservas en los M mercados de subida y bajada [28].

Para prevenir la degradación temprana del BESS, el nivel de carga se debe mantener en rangos operativos aceptables, ecuación (3.3e). Las restricciones (3.3g) y (3.3f) garantizan que el BESS cuente con energía suficiente para operar en condiciones normales y además para proveer las reservas del mercado $m = 1, \dots, M$ durante ΔT_m [28]. La restricción (3.3h) introduce el estado de carga de la batería al inicio de la etapa $t+1$. En la restricción (3.3i) se modela el estado de carga del bloque j a partir del estado de carga en $j - 1$.

Para hacer más eficiente el modelamiento del problema, en lugar de propagar directamente el vector de reservas \mathbf{q} que contiene el compromiso de reservas de la primera etapa, desde la primera etapa hacia las consecutivas, se opta por propagar los siguientes términos:

- $\sum_{m=1}^M q_{t,m}^R$ con $t \in 1, \dots, T$: ecuación (3.3c).
- $\sum_{m=1}^M q_{t,m}^L$ con $t \in 1, \dots, T$: ecuación (3.3d)
- $\sum_{m=1}^M q_{t,m}^L \cdot \Delta T_m$ con $t \in 1, \dots, T$: ecuación (3.3f))
- $\sum_{m=1}^M q_{t,m}^R \cdot \Delta T_m$ con $t \in 1, \dots, T$: ecuación (3.3g).

De esta manera, en lugar de propagar $2MT$ variables de estados para modelar los compromisos de reservas (M mercados de subida, M mercados de bajada y T etapas),

se propagan solo $4T$. Como el mercado eléctrico nacional australiano, que será el caso de estudio, tiene 3 mercados de subida y 3 de bajada ($M = 3$), se reduce en un tercio la cantidad de variables de estado requeridas para representar las reservas. Además, en pruebas preliminares, esta reparametrización de variables mostró acelerar la convergencia en la resolución del modelo. Para no perder generalidad en el modelamiento, el resto del planteamiento del problema de optimización se hace asumiendo que se propaga el vector \mathbf{q} a lo largo de las etapas.

$$\mathcal{F}_t(\mathbf{q}, s_t) = \left\{ \begin{array}{ll} P_j^{d,\text{aux}} \leq \bar{P}^d & \forall j \in (t+1, \dots, T) \quad (3.3a) \\ P_j^{c,\text{aux}} \leq \bar{P}^c & \forall j \in (t+1, \dots, T) \quad (3.3b) \\ P_j^{d,\text{aux}} - P_j^{c,\text{aux}} + \sum_{m=1}^M q_{j,m}^R \leq \bar{P}^d & \forall j \in (t+1, \dots, T) \quad (3.3c) \\ P_j^{c,\text{aux}} - P_j^{d,\text{aux}} + \sum_{m=1}^M q_{j,m}^L \leq \bar{P}^c & \forall j \in (t+1, \dots, T) \quad (3.3d) \\ \underline{S} \leq s_j^{\text{aux}} \leq \bar{S} & \forall j \in (t+1, \dots, T) \quad (3.3e) \\ s_{j-1}^{\text{aux}} + \left(P_j^{c,\text{aux}} \cdot \Delta T + \sum_{m=1}^M q_{j,m}^L \cdot \Delta T_m \right) \cdot \eta^c \leq \bar{S} & \forall j \in (t+1, \dots, T) \quad (3.3f) \\ s_{j-1}^{\text{aux}} - \left(P_j^{d,\text{aux}} \cdot \Delta T + \sum_{m=1}^M q_{j,m}^R \cdot \Delta T_m \right) / \eta^d \geq \underline{S} & \forall j \in (t+1, \dots, T) \quad (3.3g) \\ s_t^{\text{aux}} = s_t & (3.3h) \\ s_{j-1}^{\text{aux}} + \left(P_j^{c,\text{aux}} \cdot \eta^c - P_j^{d,\text{aux}} / \eta^d \right) \cdot \Delta T = s_j^{\text{aux}} & \forall j \in (t+1, \dots, T) \quad (3.3i) \\ P_{t+1}^{c,\text{aux}}, \dots, P_T^{c,\text{aux}} \in \mathbb{R}_+ & (3.3j) \\ P_{t+1}^{d,\text{aux}}, \dots, P_T^{d,\text{aux}} \in \mathbb{R}_+ & (3.3k) \\ s_t^{\text{aux}}, \dots, s_T^{\text{aux}} \in \mathbb{R}_+ & (3.3l) \end{array} \right.$$

3.3.3. Restricción de operación

En la ecuación (3.4) se define el conjunto de restricciones $\mathcal{O}_t(\mathbf{q}, s_{t-1}, P_t^d, P_t^c, s_t)$ que modelan la operación del BESS durante la etapa t , considerando la decisión de reservas \mathbf{q} y el estado de carga previo s_{t-1} . Las restricciones (3.4a) y (3.4b) representan los límites técnicos para la potencia de carga y descarga. En caso de contingencia, estos límites quedan definidos por las restricciones (3.4c) y (3.4d). Por su parte, la restricción (3.4e) garantiza que la carga del equipo se mantenga en rangos aceptables. Las restricciones (3.4g) y (3.4f) aseguran que el BESS cuente con energía suficiente para operar en condiciones normales y durante contingencias. Finalmente, la restricción (3.4h) modela el estado de carga de la etapa siguiente a partir de la decisión de inyección y el estado de carga previo.

$$\mathcal{O}_t(\mathbf{q}, s_{t-1}, s_t, P_t^d, P_t^c) = \left\{ \begin{array}{l} P_t^d \leq \bar{P}^d \quad (3.4a) \\ P_t^c \leq \bar{P}^c \quad (3.4b) \\ P_t^d - P_t^c + \sum_{m=1}^M q_{t,m}^R \leq \bar{P}^d \quad (3.4c) \\ P_t^c - P_t^d + \sum_{m=1}^M q_{t,m}^L \leq \bar{P}^c \quad (3.4d) \\ \underline{S} \leq s_t \leq \bar{S} \quad (3.4e) \\ s_{t-1} + \left(P_t^c \cdot \Delta T + \sum_{m=1}^M q_{t,m}^L \cdot \Delta T_m \right) \cdot \eta^c \leq \bar{S} \quad (3.4f) \\ s_{t-1} - \left(P_t^d \cdot \Delta T + \sum_{m=1}^M q_{t,m}^R \cdot \Delta T_m \right) / \eta^d \geq \underline{S} \quad (3.4g) \\ s_{t-1} + \left(P_t^c \cdot \eta^c - P_t^d / \eta^d \right) \cdot \Delta T = s_t \quad (3.4h) \end{array} \right.$$

3.4. Estimación estadística de la incertidumbre de precios

En la práctica, la incertidumbre de los precios de la energía tiene una distribución de probabilidades continua. Sin embargo, actualmente no existen algoritmos de resolución para problemas multi-etapa con incertidumbres de soporte continuo. En cambio, MC-SDDP requiere un modelado discreto de la incertidumbre [40, 41]. Esta sección describe el procedimiento para preprocesar datos de precios reales y obtener una representación estadística discreta de la incertidumbre.

3.4.1. Desestacionalización de datos

Para representar adecuadamente la distribución de probabilidades de los precios de la energía, es crucial atenuar la estacionalidad inherente a los datos. De no hacerlo, y al tratar con datos de distribuciones disímiles (por ejemplo, datos de distintos días de la semana o meses), la representación de la distribución de probabilidades no reflejaría fielmente el comportamiento real de los precios. Esto se debe a que datos con diferentes distribuciones se agruparían bajo una misma representación.

Para mitigar la estacionalidad de los datos, los precios de cada mes se normalizan, dividiéndolos por su promedio correspondiente, y posteriormente escalándolos por el promedio global (proceso conocido como desestacionalización). Como resultado, los promedios mensuales de los datos transformados son iguales al promedio global original. Aunque esta transformación puede parecer simple, es eficaz para reducir las diferencias entre las distribuciones de distintos meses.

3.4.2. Representación estadística de la incertidumbre

En esta investigación, la distribución de los precios de la energía se estima estadísticamente, calculando un vector de probabilidades en el caso de la primera etapa de operación y una matriz de transición en las etapas posteriores. Este cálculo se basa en un conjunto de N curvas de 24 horas del precio de la energía, representado como $\Pi = \{\vec{\pi}_i \mid i \in (1, \dots, N)\}$, donde $\vec{\pi}_i$ denota la i -ésima curva de precios. El proceso de cálculo se desarrolla de la siguiente manera:

1. **Muestreo de precios:** Se calculan los representantes de los precios observados en cada etapa t , μ_t^i con $i \in (1, \dots, R)$. Estos representantes son los centroides obtenidos al usar K-Means sobre los precios observados. Usar K-means minimiza el error cuadrático entre los representantes y los datos observados, maximizando la representatividad que tiene la discretización [41].
2. **Representación de datos observados:** Se determinan las curvas donde el precio en la etapa t queda representado por μ_t^j . Para esto, se aplica un criterio de representación por mínima distancia,

$$\mathcal{R}_t(j) = \{\vec{\pi} : \|\vec{\pi}[t] - \mu_t^j\| \leq \|\vec{\pi}[t] - \mu_t^i\|, i = 1, \dots, R, \vec{\pi} \in \Pi\} \quad (3.5)$$

3. **Estimación de probabilidades:** Para el caso de la primera etapa, la probabilidad que el precio sea μ_1^j corresponde a la proporción de veces que los precios observados en la primera etapa son representados por μ_1^j ,

$$\mathbb{P}_j^{(1)} = \mathbb{P}(\pi_1 = \mu_1^j) = \frac{|\mathcal{R}_1(j)|}{|\Pi|} \quad (3.6)$$

Para cada etapa $t > 1$, cada elemento ij de la matriz de transición $\mathbb{P}^{(t)}$ corresponde a la proporción de veces que un precio representado por μ_{t-1}^i transiciona a un precio representado por μ_t^j ,

$$\mathbb{P}_{ij}^{(t)} = \mathbb{P}(\pi_t = \mu_t^j \mid \pi_{t-1} = \mu_{t-1}^i) = \frac{|\mathcal{R}_{t-1}(i) \cap \mathcal{R}_t(j)|}{|\mathcal{R}_{t-1}(i)|} \quad (3.7)$$

3.5. Algoritmo de resolución

Aún con una representación discreta de la incertidumbre, el problema multi-etapa propuesto es computacionalmente desafiante dado que ramifica los escenarios en cada etapa. Incluso para una baja cantidad de escenarios, el árbol de escenario crece rápidamente pues depende exponencialmente del número de etapas. Para resolver este tipo de problemas se suele utilizar SDDP. Este enfoque permite aproximar la función de costos futuros

mediante cortes de Benders. En esta sección se realiza una revisión del algoritmo en su versión neutral al riesgo y con incertidumbre markoviana².

3.5.1. Formulación del problema multi-etapa

La formulación del algoritmo se lleva a cabo para un problema de minimización³ de T etapas, denotadas $t = 1, \dots, T$. En todas las etapas, excepto en la primera, existe una variable aleatoria discreta $\omega_{tj} \in \Omega_t$, donde j indica un escenario en particular. En la ecuación (3.8) se presenta el problema de primera etapa. Aquí, $x_1 \in \mathbb{R}^n$ representa la decisión, mientras que $c_1 \in \mathbb{R}^n$, $A_1 \in \mathbb{R}^{m \times n}$ y $b_1 \in \mathbb{R}^m$. La expresión $Q_{2j}(x_1, \omega_{2j})$ corresponde al costo de la segunda etapa al tomar la decisión x_1 en el escenario j .

$$\begin{aligned} z = \min_{x_1} \quad & c_1^\top \cdot x_1 + \sum_{j=1}^{|\Omega_2|} \mathbb{P}_j^{(1)} \cdot Q_{2j}(x_1, \omega_{2j}) \\ \text{s.a.} \quad & A_1 \cdot x_1 = b_1 \\ & x_1 \geq 0 \end{aligned} \tag{3.8}$$

En la ecuación (3.9) se presenta el problema de optimización de las etapas consecutivas para el escenario i . A diferencia de la primera etapa, en este caso los parámetros del problema dependen de la realización de la variable aleatoria $\omega_{ti} \in \Omega_t$. Así, $A_t(\omega_{ti}) \in \mathbb{R}^{m \times n}$, $E_t(\omega_{ti}) \in \mathbb{R}^{m \times n}$, $b_t(\omega_{ti}) \in \mathbb{R}^m$ y $c_t(\omega_{ti}) \in \mathbb{R}^n$. La variable $\lambda_t(\omega_{ti})$ corresponde al dual de las restricciones del problema. Adicionalmente, para la última etapa se impone que $Q_{T+1,j}(x_T, \omega_{T+1,j}) = 0$.

$$\begin{aligned} Q_{ti}(x_{t-1}, \omega_{ti}) = \min_{x_t} \quad & c_t(\omega_{ti})^\top \cdot x_t + \sum_{j=1}^{|\Omega_{t+1}|} \mathbb{P}_{ij}^{(t)} \cdot Q_{t+1,j}(x_t, \omega_{t+1,j}) \\ \text{s.a.} \quad & A_t(\omega_{ti}) \cdot x_t + E_t(\omega_{ti}) \cdot x_{t-1} = b_t(\omega_{ti}), \quad [\lambda_t(\omega_{ti})] \\ & x_t \geq 0 \end{aligned} \tag{3.9}$$

3.5.2. Formulación del algoritmo MC-SDDP

La formulación del algoritmo MC-SDDP se desarrolla a partir del problema multi-etapa planteado en las ecuaciones (3.8) y (3.9). Para el problema de optimización de cada etapa, la función de costos futuros Q_{tj} se sustituye por la variable de decisión θ_{tj} . Al igual que con la versión original del algoritmo SDDP, la finalidad de MC-SDDP es añadir

²El algoritmo utilizado en este trabajo se basa en la formulación propuesta en [40].

³Seguindo la tendencia observada en la literatura, en este trabajo se formula el algoritmo SDDP considerando un problema de minimización.

cortes de Benders para mejorar la aproximación de los costos futuros en cada etapa. En las ecuaciones (3.10) y (3.11) se presenta el problema de primera etapa y el de las subsiguientes, tras incorporar las variables de costos futuros y los cortes de Benders para K iteraciones del algoritmo.

$$\begin{aligned}
z_l = \min_{x_1} \quad & c_1^\top \cdot x_1 + \sum_{i=1}^{|\Omega_2|} \mathbb{P}_i^{(1)} \cdot \theta_{2i} \\
\text{s.a.} \quad & A_1 \cdot x_1 = b_1 \\
\theta_{2s} \geq -\lambda_{2,s,k}^\top \cdot E_2(\omega_{t+1,s}) \cdot (x_1 - \bar{x}_1^k) + \bar{\theta}_{2,s,k} \quad & \begin{matrix} s = 1, \dots, |\Omega_2| \\ k = 1, \dots, K \end{matrix} \\
x_1, \theta_2 \geq 0
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
\tilde{Q}_{ti}(x_{t-1}, \omega_{ti}) = \min_{x_t} \quad & c_t(\omega_{ti})^\top \cdot x_t + \sum_{j=1}^{|\Omega_{t+1}|} \mathbb{P}_{ij}^{(t)} \cdot \theta_{t+1,j} \\
\text{s.a.} \quad & A_t(\omega_{ti}) \cdot x_t + E_t(\omega_{ti}) \cdot x_{t-1} = b_t(\omega_{ti}) \quad [\lambda_t(\omega_{ti})] \\
\theta_{t+1,s} \geq -\lambda_{t+1,s,k}^\top \cdot E_{t+1}(\omega_{t+1,s}) \cdot (x_t - \bar{x}_t^k) + \bar{\theta}_{t+1,s,k} \quad & \begin{matrix} s = 1, \dots, |\Omega_{t+1}| \\ k = 1, \dots, K \end{matrix} \\
x_t, \theta_{t+1} \geq 0
\end{aligned} \tag{3.11}$$

Cada iteración del algoritmo se divide en una fase forward y una fase backward. Durante la k -ésima fase forward, comenzando desde la primera etapa, se simula la realización de la variable aleatoria, se toma la decisión de operación y se almacena en la variable \bar{x}_t^k . Mientras que en la k -ésima fase backward, comenzando desde la última etapa, se resuelve el problema $\tilde{Q}_{ts}(\bar{x}_{t-1}, \omega_{ts})$ para cada escenario $s = 1, \dots, |\Omega_t|$, y se almacenan las variable duales en $\lambda_{t,s,k}$ y el costo objetivo en $\bar{\theta}_{t,s,k}$. El pseudocódigo del procedimiento se presenta en el algoritmo 1.

Habitualmente, al término de cada fase forward, se evalúa un criterio de convergencia; si este es satisfecho, el algoritmo se detiene, si no, se sigue iterando. El criterio de convergencia se satisface cuando z_l , el límite inferior del costo total esperado (también conocido como cota inferior), se encuentra estadísticamente cerca de una estimación del costo total esperado z_u (o cota superior). En el marco de esta investigación, para determinar la cota superior z_u , se generan N escenarios aleatorios a partir del proceso estocástico. Luego, para cada escenario, se aplica la política definida por los cortes de Benders para determinar el costo incurrido, denotado como C_s con $s = 1, \dots, N$. Posteriormente, z_u corresponde al límite superior del intervalo de confianza de 95% al calcular el costo promedio $\bar{C} = \frac{1}{N} \sum_{s=1}^N C_s$.

Debido al elevado costo computacional que implica calcular la cota superior, en esta investigación no se verifica el criterio de convergencia en cada iteración, sino que se completan 100 iteraciones antes de detener el algoritmo. Al finalizar las iteraciones, para evaluar la convergencia, se calcula el gap entre z_l y z_u ,

$$\text{gap} = \frac{z_u - z_l}{|z_u|} \quad (3.12)$$

Algoritmo 1: MC-SDDP

```

1  $j = \text{None}$  // Almacena el índice del escenario previo.
2
3 for  $k = 1, \dots, K$  do // Iteraciones del algoritmo
4
5     // Fase forward
6     Se resuelve el problema de la primera etapa, ecuación (3.10).  $\rightarrow \bar{x}_1$ 
7     for  $t = 2, \dots, T$  do // Se resuelve el problema de cada etapa.
8         Se simula el proceso estocástico para determinar el escenario actual,
9         denotado  $i$ . Para  $t = 2$ , el proceso es independiente, mientras que para
10         $t \geq 3$ , es markoviano y depende del escenario previo  $j$ .  $\rightarrow i$ 
11        Se resuelve el problema  $\tilde{Q}_{ti}(\bar{x}_{t-1}, \omega_{ti})$ , ecuación (3.11).  $\rightarrow \bar{x}_t$ 
12         $j \leftarrow i$ 
13    end-for
14
15    // Fase backward
16    for  $t = T, \dots, 2$  do // Se resuelve el problema de cada etapa.
17        for  $s = 1, \dots, |\Omega_t|$  do // Se resuelve para cada posible escenario.
18            Se resuelve el problema  $\tilde{Q}_{ts}(\bar{x}_{t-1}, \omega_{ts})$ .  $\rightarrow \bar{\theta}_{t,s,k}, \lambda_{t,s,k}$ 
19            Se construye el corte de Benders utilizando  $\bar{\theta}_{t,s,k}$  y  $\lambda_{t,s,k}$ , y se agrega a
20            cada problema de la etapa  $t - 1$ . Aunque existe un problema para
21            cada escenario, todos comparten los mismos cortes como se puede
22            observar en la ecuación (3.11).
23        end-for
24    end-for
25 end-for

```

3.5.3. Simplificación en la formulación del algoritmo SDDP frente a independencias temporales

La diferencia entre el SDDP markoviano y versión que asume independencias temporales en la incertidumbre radica en que la primera requiere descomponer la función de costos futuros por escenarios y ponderarlas por la probabilidad de ocurrencia del escenario, es decir,

$$\sum_{j=1}^{|\Omega_{t+1}|} \mathbb{P}_{ij}^{(t)} \cdot \theta_{t+1,j}. \quad (3.13)$$

El algoritmo SDDP clásico asume independencia temporal en la incertidumbre, lo que significa que las realizaciones de incertidumbre en cada etapa son independientes entre sí. Este supuesto simplifica considerablemente el modelo porque permite representar la función de costos futuros como una única variable de decisión. En lugar de tener una función de costos por cada escenario futuro, $\theta_{t,j}$, se puede utilizar una única variable que represente el costo esperado, θ_t . Esta simplificación reduce la cantidad de variables de decisión involucradas en la resolución. Además, permite disminuir la cantidad de cortes necesarios en el modelamiento (expresiones (3.10) y (3.11)), ya que en cada iteración se puede realizar una ponderación de los cortes de cada escenario, ponderándolos por la probabilidad correspondiente.

3.5.4. Regularización

Aunque la convergencia del algoritmo SDDP está asegurada, tanto este como sus variantes tienden a desacelerar su velocidad de convergencia a medida que aumenta la cantidad de variables de estado⁴ [43]. Por esta razón, se suelen incorporar metodologías de regularización en las implementaciones SDDP, cuyo objetivo es mejorar la forma en que el algoritmo explora el espacio de variables de estado, del que depende la función de costos futuros. Esto permite acelerar la velocidad de convergencia, alcanzando mejores gaps para una misma cantidad de iteraciones. En esta investigación se utiliza el algoritmo de regularización propuesto en [57], una variante lineal de la regularización cuadrática originalmente propuesta en [43].

3.5.5. Estrategia de toma de decisión bajo escenarios reales

MC-SDDP determina una política de toma de decisión exclusivamente para los escenarios representados por la distribución de probabilidades de cada etapa. No obstante, en la práctica, para poder cubrir todos los posibles escenarios, se recurre a la política que mejor refleje el escenario que efectivamente está ocurriendo. De este modo, asumiendo que en la etapa t el proceso estocástico toma el valor ξ_t , se aplicará la política de operación correspondiente al escenario más cercano. Esta política la define el problema de optimización,

$$\tilde{Q}_{ti}(x_{t-1}, \xi_t) \quad \text{con} \quad i = \arg \min\{\|\xi_t - \mu_{tj}\|, j = 1, \dots, R\} \quad (3.14)$$

⁴Las variables de estado de un problema de programación dinámica corresponden a las variables que describen el estado del sistema en cada etapa del proceso de decisión, capturando toda la información relevante necesaria para la toma de decisiones futuras. En el caso del modelamiento propuesto, las variables de estado son el estado de carga de la batería y las decisiones de reserva \mathbf{q} .

3.6. Entrenamiento y simulación con datos históricos

El modelo de optimización propuesto en esta investigación tiene dos funciones: (i), decidir un monto de reservas para ofertar en el mercado eléctrico; y (ii), operar en tiempo real. Para ello, es indispensable contar con los precios de las reservas durante el horizonte de evaluación (o un forecast de los mismos), la distribución de probabilidades de los precios de la energía eléctrica y los precios de la energía observados en cada una de las etapas de operación.

Dado lo anterior, para realizar pruebas que emulen la operación real del modelo en el mercado, son necesarios dos conjuntos de datos. El primero es un conjunto de entrenamiento, que permite estimar la distribución de probabilidades de los precios de la energía. El segundo es un conjunto de prueba, que contiene el forecast de los precios de las reservas y los precios spot de la energía. Dado que el forecast del precio de las reservas es exógeno al modelamiento y a la resolución mediante SDDP, se utilizan directamente los precios observados de las reservas.

En la figura 3.1 se presenta un diagrama con el procedimiento de entrenamiento y prueba del modelo, teniendo en cuenta un horizonte de 24 horas. Del conjunto de entrenamiento, se extraen curvas de 24 horas del precio de la energía y se representa la incertidumbre (considerando la existencia de 20 escenarios por etapa). A continuación, para cada curva de 24 horas de los precios de las reservas, se formula un problema multi-etapa y se resuelve mediante SDDP. Luego, también con el conjunto de prueba, se simula la operación del BESS con curvas de 24 horas del precio de la energía y se calculan las utilidades obtenidas para cada una. Finalmente, se calcula la utilidad diaria promedio de los días simulados.

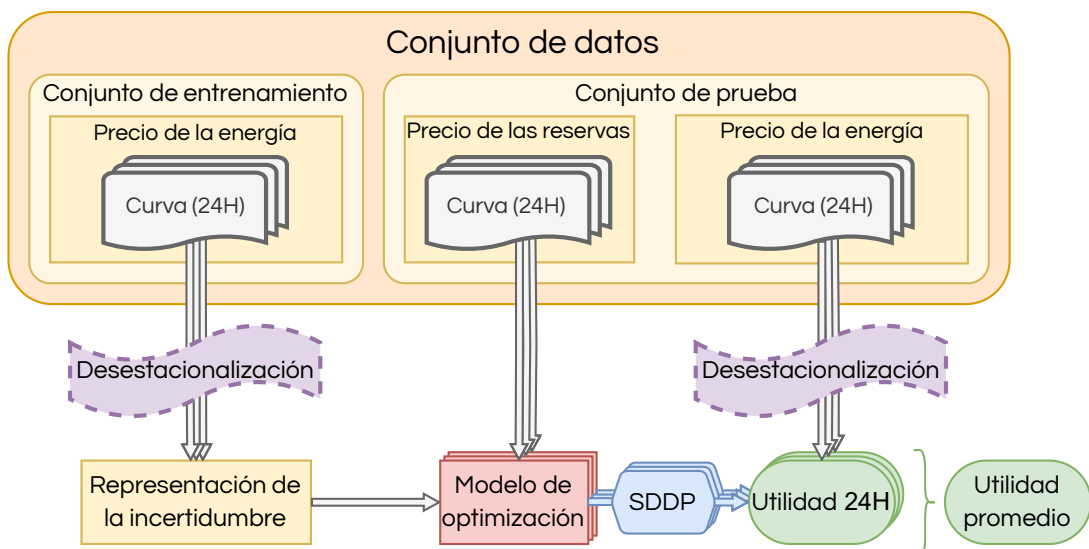


Figura 3.1: Diagrama del entrenamiento y prueba del modelo multi-etapa.

3.7. Modelos de referencia

Para evaluar las utilidades obtenidas con el modelo propuesto, éstas se comparan con las obtenidas con otros métodos de toma de decisiones, modelos basados en optimización y en aprendizaje reforzado.

3.7.1. Modelo multi-etapa independiente

Es similar al modelo propuesto, con la diferencia que la incertidumbre introducida no considera correlaciones temporales. Este tipo de modelamiento es ampliamente usado en la literatura pues permite representar un árbol de escenarios y abordar la toma dinámica de decisiones a lo largo del horizonte de evaluación. Al igual que con el modelo markoviano, dado que el problema es multi-etapa estocástico, se requiere utilizar SDDP para hacer frente al desafío de dimensionalidad y encontrar una solución en tiempos acotados.

3.7.2. Modelo determinista

A diferencia de los modelos multi-etapa, este modelo de optimización no toma decisiones dinámicas, sino que optimiza la operación basándose en la curva promedio del precio de la energía. De esta forma, se adopta una política de operación estática que, dada la curva promedio de precios de la energía, maximiza los ingresos por reservas y por arbitraje de energía. En la figura 3.2 se presenta el diagrama de entrenamiento de este enfoque. Para evaluar el rendimiento del modelo, se evalúa la calidad de la política al operar utilizando los precios de la energía del conjunto de prueba.

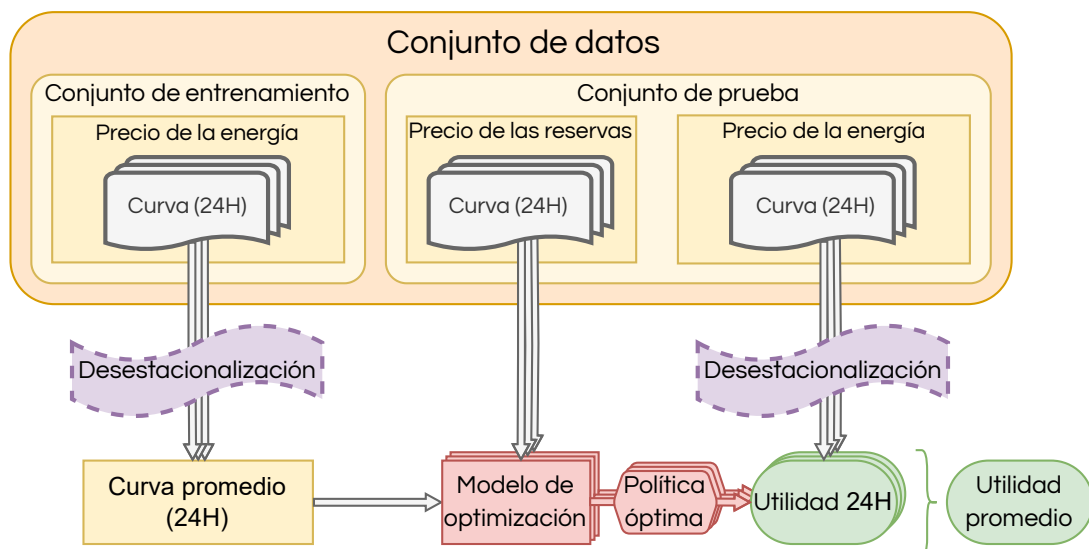


Figura 3.2: Diagrama del entrenamiento y prueba del modelo determinista.

3.7.3. Aprendizaje reforzado

En este trabajo se utiliza PPO pues permite la toma de decisiones en espacios de acción continuos, la implementación de entornos paralelos es simple y muestra una notable estabilidad durante el proceso de entrenamiento [62]. En concreto, se utiliza la implementación disponible en la biblioteca `stable-baselines3` [63].

Los hiperparámetros que componen la implementación del algoritmo, presentados en la tabla 3.1, fueron determinados manualmente por proporcionar estabilidad durante el entrenamiento. Mientras que las características utilizadas y otros hiperparámetros relacionados con el agente se pueden revisar en la Tabla 3.2. En esta última tabla, los hiperparámetros con múltiples valores se ajustaron mediante una búsqueda exhaustiva, mientras que los con valores únicos se determinaron manualmente.

Tabla 3.1: Hiperparámetros de PPO.

Hiperparámetro	Valor
Tasa de aprendizaje (<code>learning_rate</code>)	1e-3
Batch size (<code>batch_size</code>)	120
Número de transiciones simuladas antes de actualizar la política (<code>n_steps</code>)	38.400
Número de veces que se utilizan las trayectorias para actualizar el loss (<code>n_epochs</code>)	1
Rango de recorte (<code>clip_range</code>)	0,2

Tabla 3.2: Características utilizadas e hiperparámetros de la arquitectura.

Hiperparámetro	Valor
SOC	Sí
Etapa actual	Sí
Hora del día	Sí
Día de la semana	Sí, No
Mes del año	Sí, No
Precio actual	Sí
Precios antiguos (retardo)	SIN RETARDO, (2,4,6,8,10), (1,2,3,4,6,8,12)
Demanda actual	Sí
Demandas antiguas (retardo)	SIN RETARDO, (2,4,6,8,10), (1,2,3,4,6,8,12)
Dimensión espacio latente de la red del actor	16, 32, 64
Dimensión espacio latente de la red evaluadora	16, 32, 64

La arquitectura neuronal utilizada para la red del actor⁵ y la red evaluadora⁶ es idéntica y se presenta en la figura 3.3. El diseño de la arquitectura procesa las características de manera independiente, sin hacer uso de un extractor compartido. Sin embargo, antes de su procesamiento, todas las características se normalizan en un rango de -1 a 1, en

⁵Denominada “Actor Network” en inglés.

⁶Denominada “Critic Network” en inglés.

base a un rango preestablecido para cada una. Esta normalización garantiza que ninguna característica predomine durante el proceso de aprendizaje, favoreciendo la convergencia del algoritmo. En relación a las características con información cíclica como el mes, día de la semana y hora, se utiliza una codificación cíclica empleando las funciones seno y coseno.

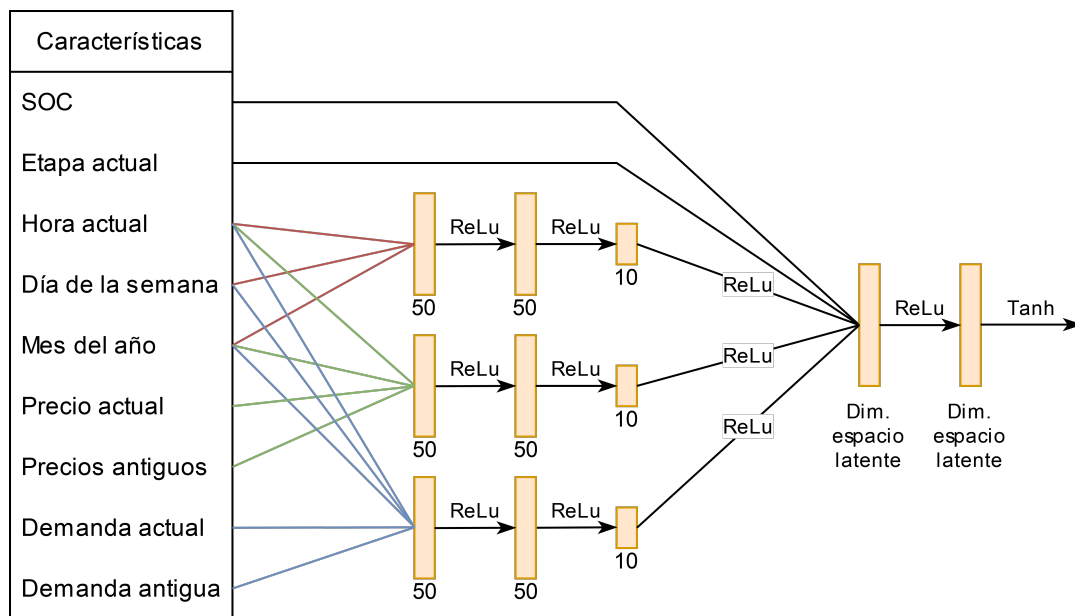


Figura 3.3: Arquitectura de la red del actor y de la red evaluadora.

Para cada combinación de hiperparámetros, se desarrollan 4 entrenamientos utilizando condiciones iniciales diferentes. Cada entrenamiento consiste en 80.000 episodios, donde cada episodio representa 24 horas de simulación con una resolución temporal de 30 minutos (que se extrae aleatoriamente del conjunto de entrenamiento). Además, para mejorar la eficiencia y estabilidad del proceso de entrenamiento, se implementa una estrategia de paralelización en la que se ejecutan simultáneamente cuatro instancias del entorno.

Finalmente, para evaluar la eficacia del ajuste de hiperparámetros, el agente se entrena nuevamente. Este entrenamiento también se realiza con cuatro condiciones iniciales diferentes y comprende 120.000 episodios con el objeto de facilitar la convergencia.

Capítulo 4

Resultados y análisis

4.1. Descripción de los casos de estudio

Para realizar un estudio exhaustivo del modelo propuesto y evaluar la ventaja potencial de utilizar modelos multi-etapa markovianos para operar sistemas BESS, se definen los siguientes casos de estudio:

- 1) **Arbitraje de energía mediante optimización (con desestacionalización, caso base):** Se compara el rendimiento del modelamiento multi-etapa markoviano con los resultados obtenidos al utilizar las otras dos técnicas de programación matemática: el modelamiento multi-etapa independiente y el modelamiento determinista. Dado que la decisión sobre reservas puede limitar la flexibilidad en la operación del BESS, para evaluar la eficacia de las decisiones dinámicas se resuelve una versión simplificada del problema de optimización propuesto. En esta simplificación, no se participa en el mercado de reservas y los ingresos provienen exclusivamente del arbitraje de energía.
- 2) **Arbitraje de energía mediante optimización (sin desestacionalización):** Con el objetivo de determinar si es necesario desestacionalizar los precios de la energía antes de estimar la distribución de probabilidades utilizada en el modelamiento multi-etapa, se realizan las mismas pruebas que en el caso anterior, pero sin desestacionalizar los precios.
- 3) **Arbitraje de energía y participación en el mercado de reservas mediante optimización:** Se aborda el problema de arbitraje de energía y optimización de reservas introducido en el capítulo anterior utilizando programación matemática. Se compara el rendimiento del modelamiento multi-etapa markoviano con los resultados obtenidos al utilizar un modelamiento multi-etapa independiente y un modelamiento determinista.
- 4) **Arbitraje de energía mediante aprendizaje reforzado:** Se emplea aprendizaje reforzado para operar el BESS, excluyendo la participación en el mercado de reservas por dos razones: 1) el algoritmo PPO es un algoritmo “sin modelo” o “model-free”, no

permitiendo la incorporación restricciones sobre las variables de decisión^{1,2}; 2) Al igual que en los casos de estudio anteriores, desde la perspectiva de evaluar la toma dinámica de decisiones, una reducción de las restricciones operativas facilita la evaluación de la política de operación.

4.2. Datos de entrada

Para cada uno de los casos de estudio, el horizonte de evaluación es de 24 horas y se utilizan los datos del mercado eléctrico nacional de Australia (NEM, por sus siglas en inglés). Este mercado suministra energía a cinco estados, a saber: Queensland (QLD), New South Wales (NSW), Victoria (VIC), South Australia (SA) y Tasmania (TAS). En cada uno de los casos de estudio, se utilizan los datos de cada estado del NEM por separado.

El NEM cuenta con seis mercados de reservas destinados a enfrentar contingencias, los cuales pueden ser de subida o bajada y operar en ventanas de 6 segundos, 60 segundos o 5 minutos. En el capítulo 5 del anexo se puede revisar en más detalle el funcionamiento de este mercado.

4.2.1. Datos históricos de precio

A continuación se describen los datos de precio y el pre-procesamiento aplicado para cada caso de estudio:

- **Arbitraje de energía mediante optimización (con y sin desestacionalización):** Los datos de entrada corresponden a los precios de la energía del año 2022. Adicionalmente, se generan curvas de precio mediante submuestreo (promediando bloques consecutivos). Con este proceso se estudia la relación entre la resolución del mercado y la utilidades recolectadas por los modelos. Se aplica validación cruzada con 5 particiones.
- **Arbitraje de energía y participación en el mercado de reservas mediante optimización:** Los datos de entrada corresponden a los precios de la energía y de las reservas ante contingencias registrados durante el año 2022. Debido al alto costo computacional del modelamiento markoviano propuesto, se evalúa el rendimiento de los modelos de optimización para una resolución de 10 minutos únicamente y solo se simula la mitad de los días del año 2022. Se aplica validación cruzada con 5 particiones.
- **Arbitraje de energía mediante aprendizaje reforzado:** Para emplear aprendizaje reforzado, se utilizan los precios de la energía de los años 2020, 2021 y 2022

¹PPO únicamente permite restringir el rango de las variables a un intervalo continuo predefinido, sin posibilidad de introducir restricciones específicas entre las acciones (i.e., variables de decisión).

²Como se mencionó en la sección anterior, si no se modela el conjunto factible de reservas en la primera etapa, las reservas comprometidas podrían no ser factibles de proveer durante la operación.

como datos de entrada. Los conjuntos de datos para entrenamiento, validación y prueba corresponden a los años 2020, 2021 y 2022, respectivamente. En estas pruebas, los precios no son desestacionalizados, ya que las redes neuronales del modelo de aprendizaje reforzado tienen la potencial capacidad de identificar estacionalidades y comportamientos no lineales, haciendo innecesaria la desestacionalización. Además, dado que parte del período cubierto por los datos corresponde a un período en el que el NEM operaba con una resolución de 30 minutos, los datos de los períodos con mayor resolución fueron muestreados a intervalos de 30 minutos³.

En los casos en que se estudian modelos basados en optimización, se excluyen los días de datos que presentan precios de la energía superiores 600[AUD/MWh] o inferiores a -600 [AUD/MWh]. Este pre-procesamiento se realiza antes del submuestreo y de desestacionalizar los datos. Por el contrario, al entrenar los modelos basados en aprendizaje reforzado no se realiza tal filtrado; sin embargo, al evaluar la operación y calcular utilidades, se excluyen los mismos días.

4.2.2. Parámetros del BESS

En la tabla 4.1 se presentan los parámetros técnicos del BESS utilizado en las pruebas. Los parámetros de eficiencia y rangos admisibles de carga corresponden a los que se suelen utilizar en la literatura [27, 29].

Tabla 4.1: Parámetros técnicos del BESS.

Parámetro	Valor
Capacidad	4 [h]
\bar{P}^c, \bar{P}^d	1 [MW]
η^c, η^d	0,95 [-]
Capacidad mínima	10 %
Capacidad máxima	90 %

4.3. Resultados

4.3.1. Comportamiento temporal de los precios de la energía

En la figura 4.1 se presentan las curvas de precio de la energía del año 2022 para el nodo ubicado en Queensland. Al desestacionalizar, se observan con mayor facilidad tendencias temporales y se reduce la dispersión intra-diaria, confirmando que esta transformación es capaz de reducir la diferencia entre curvas de precio de meses diferentes. En la figura 4.2 se presentan las matrices de correlación de las curvas de precio de la energía de los

³Desde octubre de 2021, los mercados del NEM comenzaron a operar con una resolución de 5 minutos.

años 2020, 2021 y 2022. Se observa un aumento en la correlación para los últimos años, justificando el estudio de modelos de optimización que sean capaces de reflejar procesos estocásticos con correlaciones temporales.

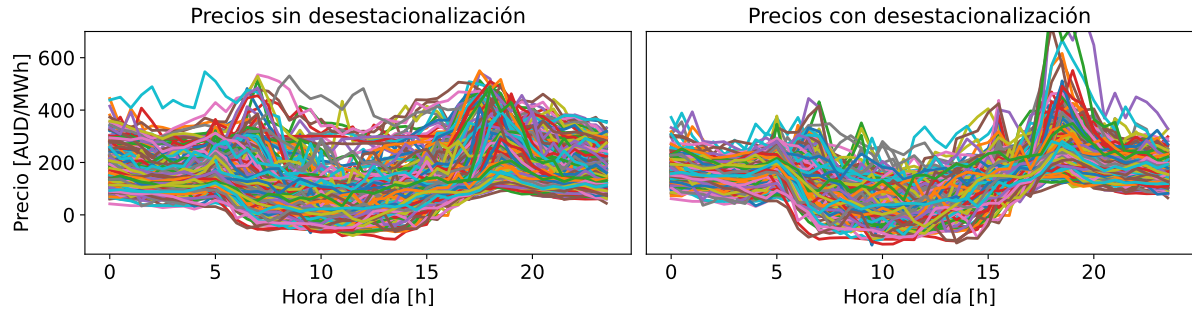


Figura 4.1: Curvas diarias del precio de la energía correspondientes al bus QLD y al año 2022. Los datos son submuestreados a una resolución de 30 minutos.

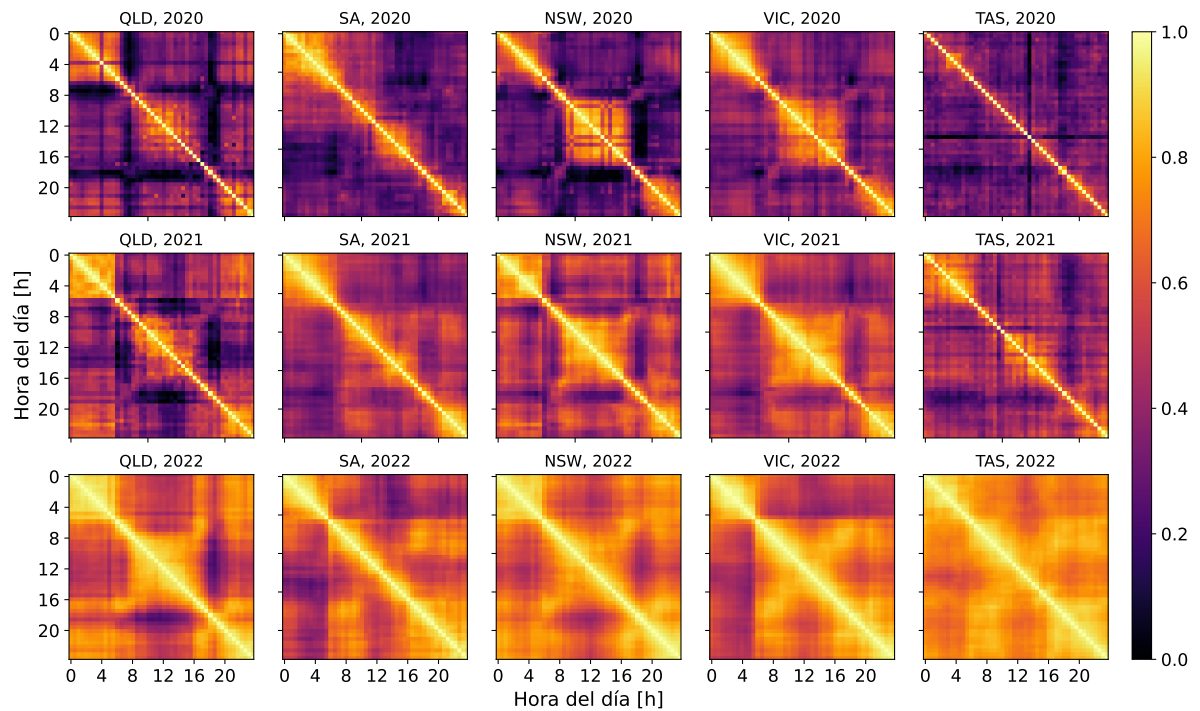


Figura 4.2: Modulo de la matriz de correlación por bus y año. Los datos son submuestreados a una resolución de 30 minutos.

4.3.2. Arbitraje de energía mediante optimización (caso base)

La figura 4.3 muestra las utilidades de los modelos basados en optimización al arbitrar energía, relativas al modelo markoviano. Se observa que el modelo markoviano supera en utilidades al modelo multi-etapa independiente entre un 5% y un 12%, y al modelo

determinista entre un 9% y un 45%. Esto indica que, al arbitrar energía con un BESS bajo incertidumbre de precios, incorporar las correlaciones temporales en la representación de la incertidumbre, en lugar de suponer independencia, resulta en mejores políticas de operación. Aunque es subóptimo modelar la incertidumbre asumiendo independencias temporales, genera mejores políticas de operación que una política estática, óptima para el promedio de casos.

En arbitraje de energía, cuando se utiliza una resolución de 5 minutos, el costo computacional para entrenar el modelamiento multi-etapa markoviano es del orden de 4 minutos, con un uso de memoria de aproximadamente 2,5 GB. En comparación, el modelamiento multi-etapa independiente requiere aproximadamente 2 minutos y 0,25 GB, mientras que el modelo determinista se resuelve en cuestión de segundos con un uso de 0,15 GB (como se muestra en la tabla B.1 del anexo). Aunque el modelamiento multi-etapa markoviano demanda más recursos computacionales (memoria y tiempo de computo), los mayores ingresos que permite alcanzar justifican, preliminarmente, su estudio.

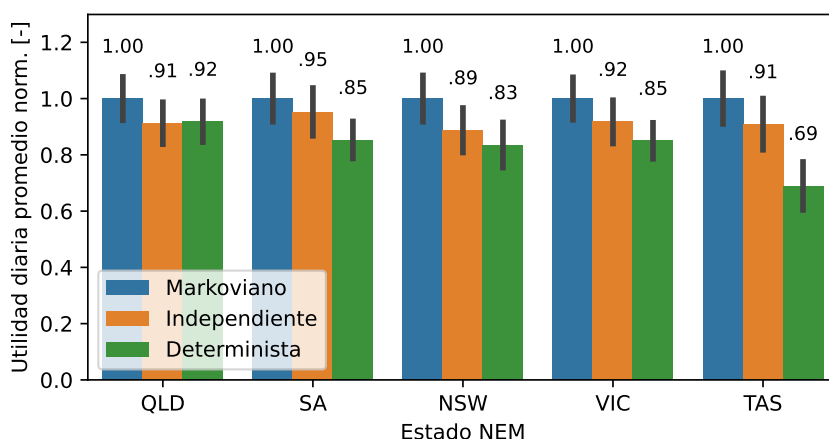


Figura 4.3: Utilidades relativas de los modelos de optimización al arbitrar energía y desestacionalizar precios, mercado con resolución de 5 minutos. El error de estimación de la media considera un 95% de confianza.

La figura 4.4 presenta las utilidades obtenidas al variar la resolución con la que opera el mercado, relativas al modelo markoviano con resolución de 5 minutos. Se constata que los ingresos de cada modelo son menores a medida que disminuye la resolución del mercado. Este resultado es relevante pues, debido a que los BESS tienen una velocidad de respuesta rápida, una baja resolución en la operación del mercado limita la toma de decisiones, reduce los ingresos y extiende el plazo de recuperación de las inversiones. Respecto a las utilidades de cada modelo, el modelo determinista es el menos beneficiado con un aumento en la resolución del mercado, mientras que el modelo independiente es el que más se beneficia.

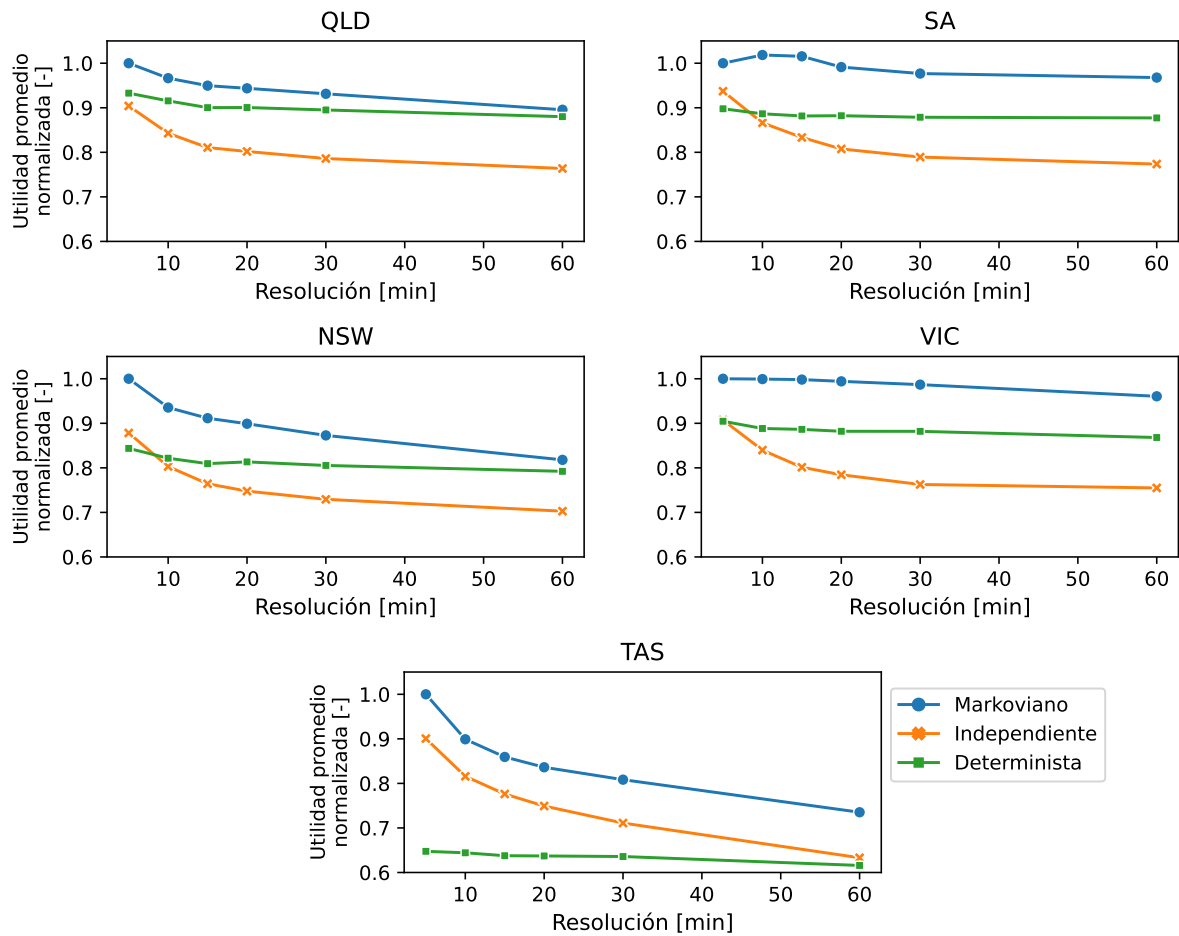


Figura 4.4: Utilidades relativas de los modelos de optimización al arbitrar energía. Se desestacionalizan los precios y se consideran diferentes resoluciones en la operación del mercado.

4.3.3. Arbitraje de energía mediante optimización (sin desestacionalización)

La figura 4.5 muestra las utilidades de los modelos basados en optimización al arbitrar energía, tanto si se desestacionalizan los precios de la energía como si no se desestacionalizan, relativas al modelo markoviano en que se aplica desestacionalización. En los casos en que se desestacionalizan los precios, las utilidades se reestacionalizan usando los mismos factores aplicados en la desestacionalización. Este proceso de reestacionalización corrige las utilidades y permite una comparación directa con las utilidades del caso en que no se desestacionaliza.

Cuando se desestacionalizan los precios, los modelos multi-etapa logran utilidades superiores en comparación con el caso en que no se desestacionalizan. En la escala relativa utilizada, se observan incrementos de entre 0,04 y 0,15 puntos para el caso markoviano y entre 0,09 y 0,23 puntos para el caso independiente. Esto respalda la idea de que, para

una representación multi-etapa, se debe maximizar la similitud entre las curvas de precio utilizadas para estimar la distribución, tal como se discutió en la Subsección 3.4.1. Por el contrario, el modelo determinista muestra una mínima variación en las utilidades, ya que identifica correctamente cuándo comprar o vender energía al operar basado en la curva de precio promedio, ya sea con o sin desestacionalización.

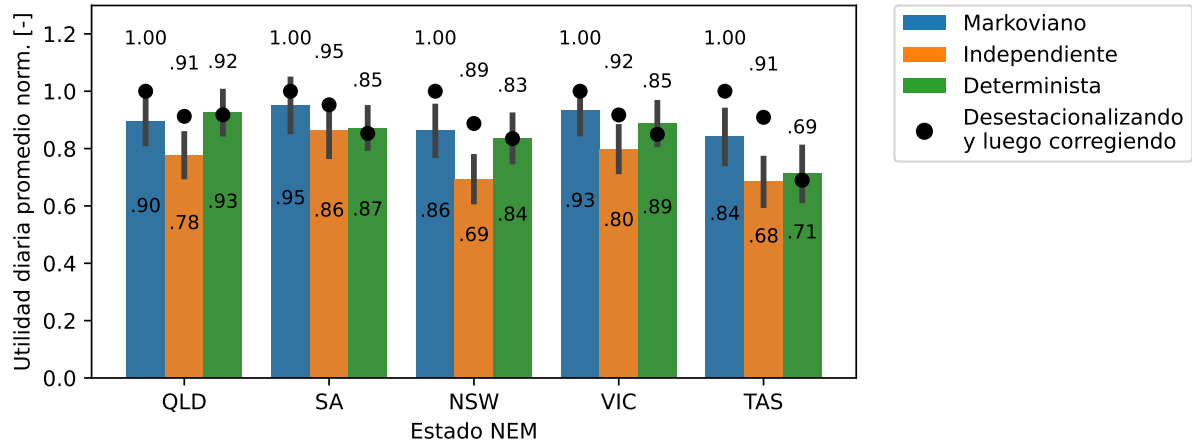


Figura 4.5: Utilidades relativas de los modelos basados en optimización al arbitrar energía, mercado con resolución de 5 minutos. Las barras representan las utilidades al no desestacionalizar precios. Los puntos representan las utilidades al desestacionalizar los precios de la energía y luego corregir las utilidades.

El anexo contiene detalles sobre las utilidades relativas al no desestacionalizar los precios y utilizar diferentes resoluciones en la operación del mercado, figura B.1. En la figura B.2 del anexo se muestran las utilidades promedio para los modelos basados en optimización al arbitrar energía, en escala absoluta, tanto si se desestacionalizan y como si no se desestacionalizan los precios de la energía.

4.3.4. Arbitraje de energía y participación en el mercado de reservas mediante optimización

En la figura 4.6 se presentan las utilidades obtenidas al usar el modelo propuesto en la metodología, relativas al modelo markoviano. Se observa que el modelo determinista iguala o supera marginalmente al modelo markoviano, mientras que el modelo markoviano supera entre 0.06 a 0.16 puntos al modelo independiente. Esto podría indicar que, en ciertos casos, la representación independiente de la incertidumbre es imprecisa y que podría ser más beneficioso utilizar un modelo estático.

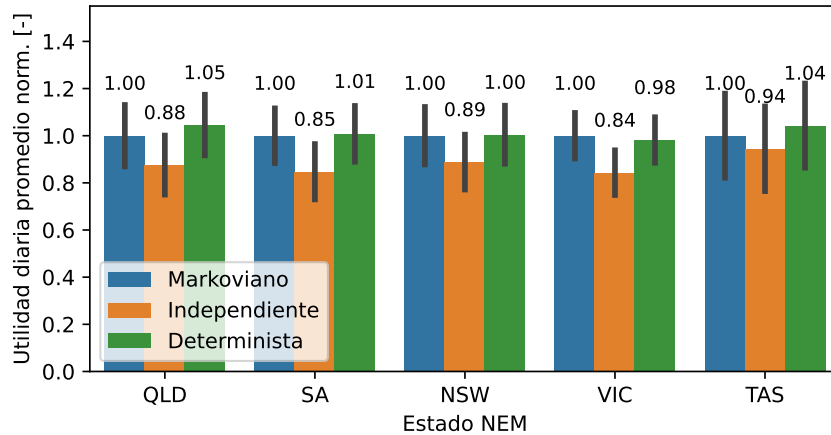


Figura 4.6: Utilidades relativas de los modelos basados en optimización al participar en el mercado de reservas y arbitrar de energía, mercado con resolución de 10 minutos y precios desestacionalizados. Se excluyen los días con utilidades superiores a 5.000[AUD].

Es probable que el modelo determinista iguale o supere al modelo markoviano por dos razones. Primero, al modelar la participación en el mercado de reservas, la convergencia de los modelos multi-etapa empeora, provocando que la solución encontrada en la primera etapa subestime costos de operación y que la operación dinámica sea subóptima. Segundo, el compromiso de reservas reduce el margen para operar el BESS, disminuyendo la importancia de las decisiones dinámicas. En otras palabras, los posibles ingresos por reservas son comparables a los ingresos esperados por arbitraje, lo que lleva a priorizar la participación en el mercado de reservas y a que la operación de la batería siga un programa acorde con los compromisos adquiridos en este mercado.

En la tabla 4.2 se presenta el gap de convergencia promedio de los problemas de optimización multi-etapa, calculados utilizando la expresión (3.12). Para una resolución de 10 minutos, en el caso markoviano, se observan gaps entre un 14% y 19% para arbitraje y participación en el mercado de reservas, mientras que para solo arbitraje, los gaps varían entre 7% y 9%. En el caso multi-etapa independiente, los gaps promedio varían entre 7% y 11% para arbitraje y participación en el mercado de reservas, y son cercanos al 3% para solo arbitraje de energía.

Se observa que los gaps de convergencia porcentuales aumentan cerca del doble al incorporar decisiones de reservas en el modelamiento, en comparación con el caso en que solo se arbitra energía. Esto hace complejo determinar con certeza el rendimiento del modelamiento multi-etapa markoviano respecto al modelamiento determinista. Como se mencionó en la sección anterior, se realizaron 100 iteraciones para resolver cada modelo multi-etapa; aumentar el número de iteraciones podría reducir el gap de las soluciones encontradas, aunque esto requeriría una mayor cantidad de recursos computacionales (memoria y tiempo de ejecución). Este trabajo incorporó un mecanismo de regularización en la implementación del algoritmo SDDP. Sin embargo, existen metodologías de regularización más efectivas, como la utilizada en [43], que podrían reducir los gaps obtenidos en esta investigación. Además, se podría implementar una resolución paralelizada asincrónica del

algoritmo SDDP, como la propuesta en [44], para incrementar la cantidad de iteraciones por unidad de tiempo de cálculo y contener el tiempo de ejecución al elevar la cantidad de iteraciones.

Al incorporar la asignación de reservas en el modelamiento, la cantidad variables de estados de los modelos multi-etapa se incrementa de 1 a $1 + 4T$, con T siendo la cantidad de etapas⁴. Así, para una resolución de 10 minutos, la cantidad de estados aumenta de 1 a 577. El aumento en la cantidad de estados ralentiza la convergencia, ya que dificulta la aproximación de la función de costos futuros al depender de más variables, requiriendo más iteraciones para que el algoritmo SDDP converja.

Tabla 4.2: Gap promedio de los modelos de optimización multi-etapa utilizados en esta investigación.

Algoritmo	Res. [min]	Arbitraje					Arbitraje y reservas				
		QLD	SA	NSW	VIC	TAS	QLD	SA	NSW	VIC	TAS
MC-SDDP	60	-0,3%	2,1%	-1,3%	1,8%	2,6%	-	-	-	-	-
	30	11,8%	11,1%	15,2%	15,3%	16,5%	-	-	-	-	-
	20	5,5%	2,8%	3,1%	0,7%	4,8%	-	-	-	-	-
	15	9,1%	6,5%	7,2%	5,5%	5,8%	-	-	-	-	-
	10	8,7%	7,1%	8,4%	8,4%	7,8%	17,5%	14,6%	18,5%	16,8%	14,2%
	5	9,0%	9,1%	10,6%	9,1%	9,3%	-	-	-	-	-
SDDP	60	2,8%	3,6%	3,8%	3,9%	4,0%	-	-	-	-	-
	30	3,2%	2,3%	3,0%	2,7%	3,4%	-	-	-	-	-
	20	3,0%	3,7%	3,1%	3,5%	4,2%	-	-	-	-	-
	15	2,8%	2,2%	2,4%	2,3%	2,9%	-	-	-	-	-
	10	2,9%	2,5%	2,8%	2,6%	2,8%	10,7%	9,4%	10,2%	8,6%	7,3%
	5	2,9%	3,1%	2,5%	2,8%	2,8%	-	-	-	-	-

En la figura 4.7 se presentan las utilidades promedio para cada tipo de modelamiento, tanto al participar como al no participar en el mercado de reservas. Dado que los datos de precio contienen días en que ocurrieron fallas, para calcular las utilidades promedio se excluyen los días con utilidades superiores a 5000[AUD]⁵. Es posible verificar que, en todos los buses, las utilidades superan a las del caso que sólo arbitra energía. Así, a pesar que convergencia de los modelos multi-etapa es baja, se alcanzan mayores utilidades cuando el modelo sí integra la participación en los mercados de reservas.

⁴Considerando la reparametrización de la decisión de reserva propuesta en el capítulo anterior.

⁵En el gráfico B.4 del anexo sólo se excluyen días con utilidades mayores a 10000[AUD].

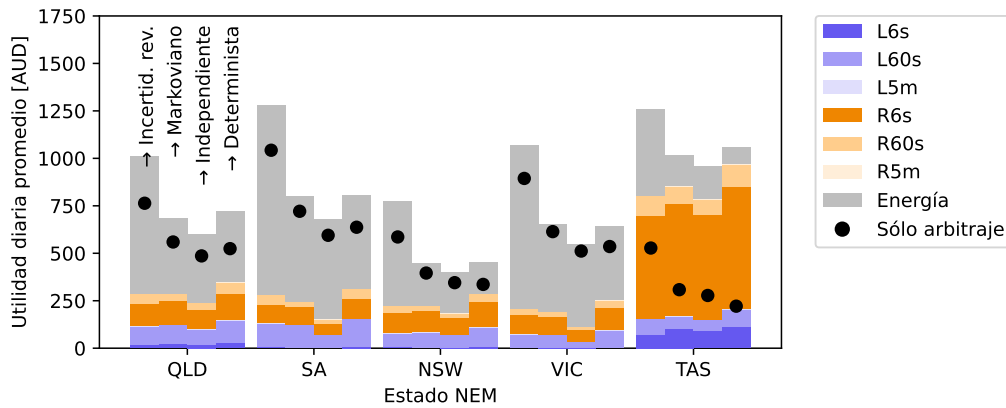


Figura 4.7: Utilidades de los modelos de optimización al participar en el mercado de reservas y arbitrar de energía, mercado con resolución de 10 minutos y precios desestacionalizados. Se excluyen los días con utilidades superiores a 5.000[AUD]. Referencialmente, se presentan las utilidades de un modelo que revela la incertidumbre en la primera etapa (primera barra de cada estado).

Los recursos computacionales requeridos para el entrenamiento de los diferentes modelamientos se presentan en la tabla B.1 del anexo. En arbitraje y participación en el mercado de reservas, para una resolución de 10 minutos, el modelo multi-etapa markoviano requiere aproximadamente 40 minutos de entrenamiento y utiliza alrededor de 13 GB de memoria. En contraste, el modelo determinista se resuelve en cuestión de segundos y utiliza cerca de 0,16 GB.

El aumento en la complejidad computacional del modelamiento multi-etapa al incluir reservas, junto con las dificultades adicionales de utilizar un modelamiento multi-etapa markoviano (como mayor uso de memoria y tiempo de cómputo), y que no escala favorablemente con la resolución del problema (no fue posible simular el modelo a una resolución de operación del mercado de 5 minutos), lleva a que una aproximación no estocástica y sin capacidad de adaptación dinámica tenga mejores resultados (política estática).

Es importante destacar que, aunque existe una relación entre los recursos computacionales y la cantidad de variables de estado manejadas en la convergencia de los algoritmos SDDP [43], los resultados e implicancias presentados en esta sección se limitan al problema de arbitraje que incorpora decisiones de reservas y únicamente cuando la fuente de incertidumbre son los precios de la energía. Resulta difícil extrapolar los resultados a otros problemas, ya que podrían depender de múltiples factores, entre los que se encuentran: el beneficio obtenido al ajustar dinámicamente la operación, cuánto influye la decisión de la primera etapa en la libertad operacional, la relación entre los ingresos de primera etapa y los posibles ingresos del resto de etapas, y la cantidad de variables de estado.

En el caso particular de arbitraje y participación en el mercado de reservas, podría ser beneficioso contar con la capacidad de ajuste dinámico en respuesta a la señal de precios de la energía a lo largo de la operación. Sin embargo, debido a la construcción del problema, la decisión de la primera etapa condicionará la flexibilidad operativa de la batería

en las etapas siguientes. Un mayor compromiso de reservas ajustará los niveles admisibles de estado de carga y potencia, mientras que un menor compromiso los relajará. Al inspeccionar la figura 4.7, se observa que los ingresos de la primera etapa por compromisos de reserva representan una parte no despreciable de los ingresos totales. Además de estos factores, que disminuyen la relevancia de los beneficios de una operación dinámica frente a una operación estática, se encuentra que la cantidad de variables de estado es 577, entre dos y tres órdenes de magnitud superior al caso en que solo se arbitra energía, dificultando la convergencia.

La relación entre los potenciales ingresos de primera etapa y los de las etapas de operación no depende de la naturaleza del problema, sino directamente de las dinámicas del mercado eléctrico del que provienen los datos. En otras palabras, para el año 2022, los mercados de reservas pudieron ofrecer más oportunidades que desarrollar estrategias de arbitraje de energía. Una discusión similar podría plantearse en torno al caso de arbitraje de energía puro y el rendimiento superior del modelo multi-etapa markoviano. En años anteriores, como 2020 o 2021, se observaron niveles inferiores de correlación en los precios de energía, lo que podría haber hecho menos conveniente utilizar una metodología multi-etapa markoviana para operar sistemas BESS (ver figura 4.2).

Esta situación resalta la importancia de elegir una metodología de toma de decisiones de manera holística. No solo se debe considerar la capacidad teórica para representar el fenómeno, sino también que la metodología permita alcanzar óptimos razonables en la práctica, considerando los recursos computacionales disponibles (tiempo de ejecución y uso de memoria) y otras variables que podrían afectar el desempeño, como las características particulares de los datos involucrados. Esto es crucial, ya que podrían existir metodologías alternativas más costo-efectivas que, con menor sofisticación y requerimientos computacionales, logran resultados comparables.

4.3.5. Arbitraje de energía mediante aprendizaje reforzado

En la tabla 4.3 se muestran los hiperparámetros ajustables óptimos. Estos hiperparámetros se ajustaron con los datos del nodo de referencia ubicado en Queensland y posteriormente fueron utilizados para entrenar agentes con los datos del resto de nodos.

Tabla 4.3: Características e hiperparámetros óptimos.

Hiperparámetro	Valor
Día de la semana	No
Mes del año	Sí
Precios antiguos (retardo)	(1,2,3,4,6,8,12)
Demandas antiguas (retardo)	(2,4,6,8,10)
Dimensión espacio latente de la red del actor	32
Dimensión espacio latente de la red evaluadora	64

La figura 4.8 presenta las utilidades relativas obtenidas al utilizar modelos basados

en optimización y en aprendizaje reforzado. Para más detalles, se pueden consultar las utilidades en escala absoluta en el anexo, en la figura B.5. Los modelos de optimización se entrenan y simulan con datos desestacionalizados, luego las utilidades se reestacionalizan. Las comparativas presentadas se realizaron con una resolución de 30 minutos, ya que, como se mencionó previamente, el período cubierto por los datos utilizados para entrenar el modelo basado en aprendizaje reforzado corresponde a un período en el que el NEM operaba con esta resolución.

Para una resolución de 30 minutos, los modelos multi-etapa markovianos obtienen resultados superiores al realizar arbitraje de energía en comparación con políticas estáticas de operación, modelamientos multi-etapa con incertidumbre sin correlaciones y modelos basados en aprendizaje reforzado. En particular, el modelamiento multi-etapa markoviano supera al modelamiento multi-etapa independiente entre un 15 % y un 27 %, al modelamiento determinista entre un 5 % y un 23 %, y a la toma de decisiones mediante aprendizaje reforzado entre un 11 % y un 32 %.

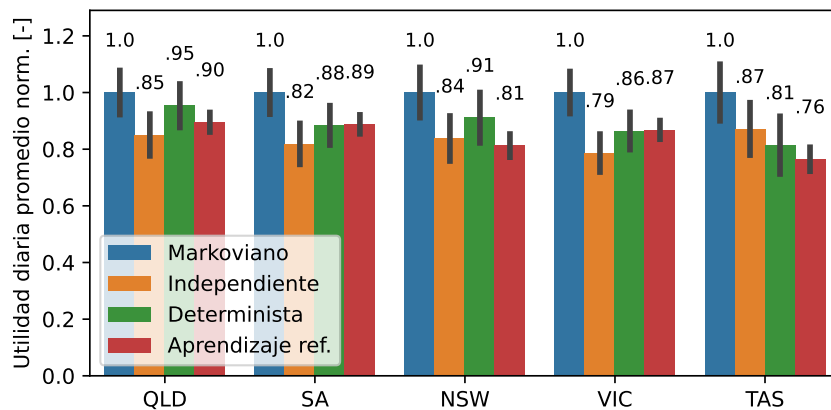


Figura 4.8: Utilidades relativas de los modelos de optimización y el basado en aprendizaje reforzado al realizar arbitraje de energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.

A pesar de que el modelo basado en aprendizaje reforzado utiliza más información para tomar decisiones que los modelos basados en optimización, su rendimiento es menor o comparable al del modelo determinista. Esto es interesante, ya que el modelo determinista no toma decisiones dinámicas, pero funciona de manera óptima para la curva promedio de precio. En cambio, el modelo de aprendizaje reforzado no resuelve un problema de optimización; en su lugar, emplea un vector de características y redes neuronales para tomar decisiones y estimar costos futuros. La falta de un modelo explícito en el aprendizaje reforzado dificulta la captura de las relaciones entre estados, acciones e ingresos, requiriendo una gran cantidad de recursos computacionales para entrenarse.

Los recursos computacionales necesarios para el entrenamiento de cada modelo se presentan en la tabla B.1 del anexo. Para una resolución de 30 minutos, el modelo basado en aprendizaje reforzado requiere aproximadamente 50 minutos y alrededor de 2,5 GB

de memoria. El modelo multi-etapa markoviano necesita alrededor de un minuto para su resolución y utiliza aproximadamente 0,5 GB de memoria. En contraste, el modelo determinista se resuelve en cuestión de segundos y requiere cerca de 0,14 GB de memoria.

Pese a los resultados obtenidos, no se debe descartar completamente al aprendizaje reforzado como mecanismo para arbitrar energía con un BESS. Una de las ventajas del uso de aprendizaje reforzado es que integra datos históricos de manera endógena, mientras que con un modelo markoviano esto debe realizarse exógenamente mediante la estimación de una matriz de transición. Además, el modelo markoviano presenta otro problema: la toma de decisiones depende únicamente del precio presente y el estado de carga. En algunos casos, este supuesto podría no ser lo suficientemente realista, considerando que los precios futuros pueden estar influenciado por precios más antiguos que el de la etapa presente o por otros factores, como la demanda. Ampliar el conjunto de entrenamiento del modelo basado en aprendizaje reforzado a más de un año de datos podría mejorar su rendimiento, ya que los modelos de machine learning identifican patrones con mayor precisión cuando se utilizan más datos para el entrenamiento.

Adicionalmente, existen herramientas para mejorar el rendimiento de modelos basados en aprendizaje reforzado que no se implementaron en este documento, como el uso de redes neuronales recurrentes [20, 64, 65]. También, se puede enriquecer el modelo al incorporar más información en la toma de decisiones del modelo. Preliminarmente, se contempló la inclusión de datos de radiación solar y temperatura, lo que resultó en un incremento en las utilidades de Queensland de entorno a un 3%.

Las figuras 4.9 y 4.10 muestran la operación del BESS al utilizar modelos basados en optimización y aprendizaje reforzado, respectivamente. Los datos empleados corresponden a las primeras dos semanas del 2022 del bus de referencia ubicado en Queensland, y se opera para un horizonte de 24 horas (cada 24 horas la batería inicia con el mínimo estado de carga posible). Según lo observado, el modelo determinista y el markoviano presentan dos ciclos de carga-descarga. El primero ocurre antes del mediodía y el segundo, se extiende desde la mañana hasta la noche. Esto concuerda con la curva de precios observada y con el comportamiento del modelo con incertidumbre revelada. En contraste, el modelo independiente presenta un comportamiento menos regular, con días de operación en los que el BESS no se carga completamente o se descarga antes del precio máximo (como se observa en los días 6, 7, 8 y 12).

En relación al modelo basado en aprendizaje reforzado, este presenta un único ciclo de carga-descarga que inicia al mediodía, cuando los precios de la energía son bajos, y finaliza después del precio máximo observado. La existencia de un único ciclo de carga por día sugiere que la política de operación podría no capturar completamente la dinámica de precios, pues el modelo markoviano, que tiene una mejor política, realiza dos ciclos de carga por día. En la figura 4.10 también se muestra la curva de demanda de Queensland. Se aprecia que los precios más altos ocurren en los momentos de mayor demanda, resaltando la importancia que el modelo no sólo considere información de precios para tomar decisiones, sino también variables que proporcionen información sobre su evolución.

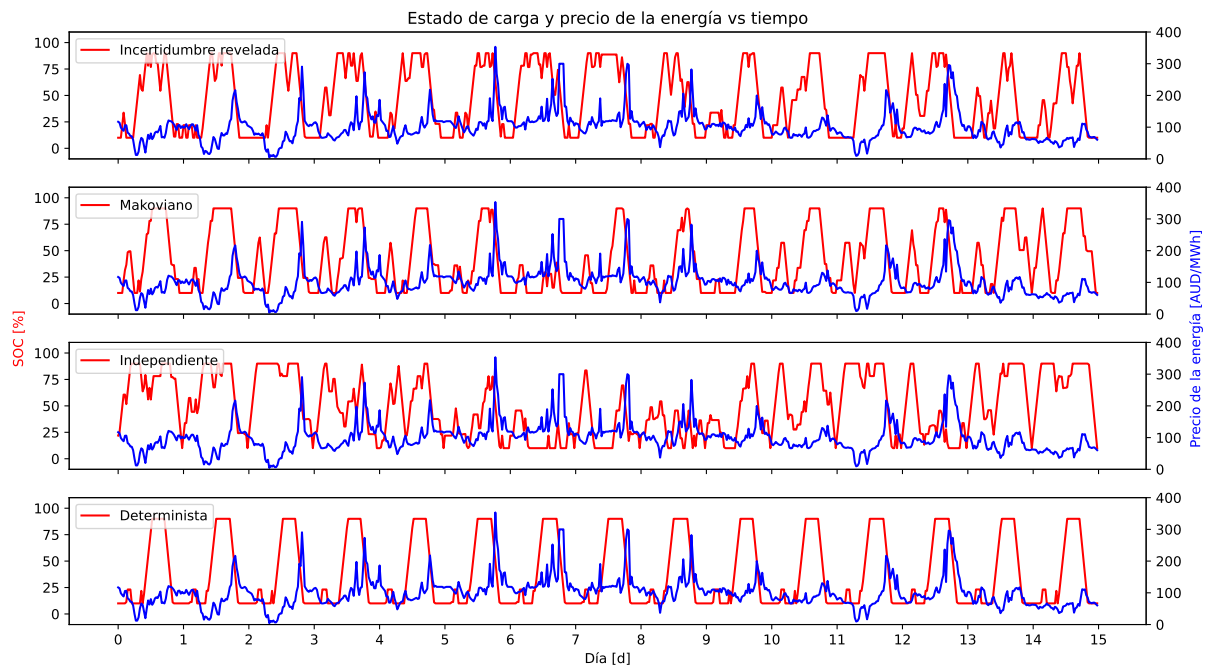


Figura 4.9: Operación de los primeros 15 días del año 2022 al utilizar modelos basados en optimización para arbitrar energía. Datos del nodo de referencia ubicado en Queensland.

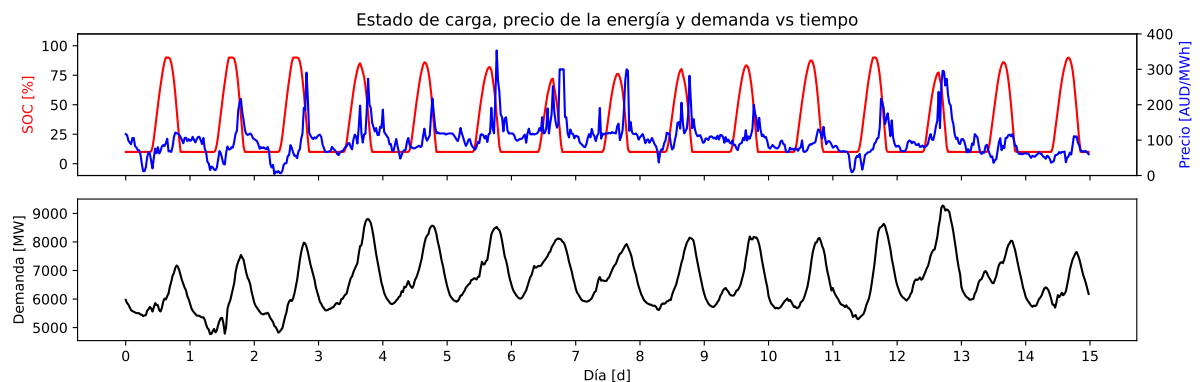


Figura 4.10: Operación de los primeros 15 días del año 2022 al utilizar aprendizaje reforzado para arbitrar energía. Datos del nodo de referencia ubicado en Queensland.

En la figura 4.11 se muestran las distribuciones de las utilidades de los modelos basados en optimización y los basados en aprendizaje reforzado. Un aspecto destacable es que el modelo basado en aprendizaje reforzado presenta colas inferiores más cortas que los modelos basados en optimización. Esta situación probablemente se deba a que el modelo basado en aprendizaje reforzado utiliza más información para tomar decisiones, lo que le podría permitir identificar días menos favorables para la operación y actuar en consecuencia. Por el contrario, los modelos de optimización, incluso el markoviano, podrían tener dificultades para reducir pérdidas debido a que no pueden identificar dinámicas que se desvían de la representación estadística de la incertidumbre.

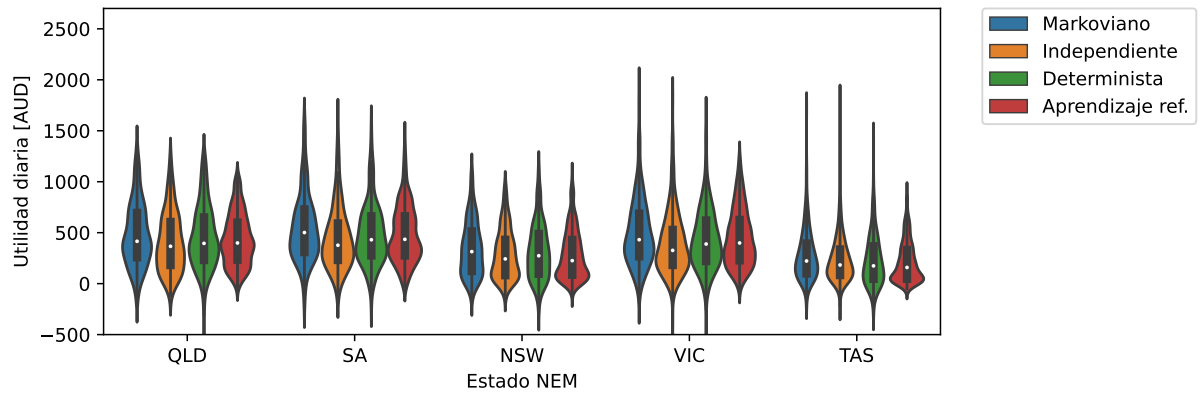


Figura 4.11: Distribución de las utilidades de los modelos basados en optimización y en aprendizaje reforzado al arbitrar energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.

Considerando los antecedentes presentados en este capítulo, los modelos multi-etapa markovianos obtienen resultados superiores al realizar arbitraje de energía en comparación con políticas estáticas de operación, modelamientos multi-etapa con incertidumbre sin correlaciones y modelos basados en aprendizaje reforzado. Sin embargo, cuando se incorpora la participación en mercados de reserva, las políticas de operación estáticas logran utilidades promedio similares a las del modelamiento markoviano. Esto probablemente se debió al bajo nivel de convergencia obtenido en los modelos multi-etapa, ocasionado por el aumento en la cantidad de variables de estado del problema.

En ambos casos de estudio, tanto al considerar como al no considerar reservas, los modelos multi-etapa markovianos obtienen mejores resultados que los modelos multi-etapa independiente. Esto demuestra que cuando la fuente de incertidumbre son los precios de la energía y se utilizan modelamientos multi-etapa para operar, adoptar representaciones que incorporan correlaciones en la incertidumbre podrían permitir alcanzar políticas más eficientes que al ignorar dichas correlaciones.

No obstante, al elegir metodologías de operación, no se debe desatender el criterio de costo-eficiencia, ya que una mayor complejidad matemática en el modelamiento no garantiza mayores rendimientos en la práctica. Por ejemplo, al arbitrar energía con aprendizaje reforzado, una metodología que considera, además de los precios de la energía, los precios y demandas en horas previas y que requirió cerca de 50 minutos de entrenamiento, no genera utilidades consistentemente mejores que una programación estática con tiempos de ejecución de segundos. Esto, nuevamente, resalta la importancia de seguir criterios más completos al seleccionar metodologías de modelamiento y operación, ya que opciones con una capacidad para representar fenómenos complejos requieren supuestos y herramientas adicionales para alcanzar rendimientos óptimos (tiempo de ejecución, memoria, datos de entrenamiento, elección de hiperparámetros, entre otros).

Capítulo 5

Conclusión

En este estudio se investigó el uso de modelos de optimización multi-etapa con representación markoviana de la incertidumbre para operar sistemas de almacenamiento. El problema de operación analizado consideró arbitraje de energía, tanto con como sin participación en mercados de reservas, con incertidumbre en los precios de la energía. En este contexto, se comparó el desempeño de la política de operación derivada del modelo multi-etapa markoviano con respecto al de otras metodologías de toma de decisión.

Luego de las pruebas realizadas, se concluye que un modelamiento markoviano de la incertidumbre mejora las utilidades en el problema de arbitraje de energía cuando los precios de la energía constituyen la fuente de incertidumbre. Al considerar la participación en el mercado de reservas, esta representación de la incertidumbre también genera mayores utilidades en comparación con el modelamiento que asume independencia temporal. No obstante, al participar en el mercado de reservas, el modelamiento determinista resulta igual o más eficaz que los modelos multi-etapa. Esta situación probablemente se debe a que al considerar reservas, la complejidad computacional del problema aumenta, empeorando la convergencia del modelo y derivando en políticas subóptimas. Debido a esto, no es posible determinar con certeza si el rendimiento del modelamiento multi-etapa markoviano es ineficaz para resolver el problema o si las dificultades de convergencia son la causa de los resultados inferiores.

En base a lo anterior, la hipótesis se verifica de manera parcial. Para el problema de arbitraje, un modelamiento multi-etapa con representación markoviana de la incertidumbre permite superar los resultados del resto de metodologías de toma de decisión. Sin embargo, cuando el problema involucra participación en el mercado de reservas, no es concluyente la superioridad del modelamiento markoviano respecto al modelo determinista. No obstante, en todos los casos de estudio, se concluye que la representación markoviana de la incertidumbre produce mayores utilidades en la operación en comparación con la representación independiente. Se debe proceder con cautela al extrapolar estas conclusiones a otros tipos de problemas, ya que los resultados obtenidos son intrínsecamente dependientes de las características del problema. Las características del fenómeno modelado (variables de decisión, función objetivo, restricciones, cantidad de variables de estado), la cantidad de etapas y las características propias de los datos (por ejemplo, ingresos por

etapa y nivel de correlaciones en la incertidumbre) son elementos que podrían impactar la generalización de los resultados.

Los resultados de esta investigación no se limitan solo a la conveniencia de modelar la incertidumbre de precios de la energía considerando correlaciones. Al elegir el enfoque de modelamiento, es crucial tener en cuenta criterios integrales, no solo la capacidad de representar fenómenos complejos, de manera de asegurar una relación costo-beneficio ventajosa en la práctica. Esto podría implicar utilizar metodologías más modestas cuando la implementación de herramientas más sofisticadas sea contraproducente en la práctica. En esta investigación en particular se observó que, cuando los recursos computacionales son limitados; existen dificultades en la convergencia que no garantizan desempeños óptimos; los datos son escasos o las ventajas del modelamiento son desaprovechadas por los datos (por ejemplo, si los precios de reservas son muy altos y se privilegian estos mercados); metodologías más sencillas podrían ser más efectivas que otras más complejas.

En el caso del aprendizaje reforzado, al considerar el problema de arbitraje de energía con resolución de 30 minutos, se observan mayores costos computacionales y menores utilidades en comparación con el modelamiento multi-etapa markoviano. Sin embargo, esta metodología cuenta con otras ventajas, como no requerir herramientas de regularización de datos, como la desestacionalización, lo que facilita su implementación práctica. Las pruebas indican que el modelo basado en aprendizaje reforzado podría reducir las colas inferiores en las utilidades percibidas, disminuyendo el riesgo al arbitrar energía. Futuras extensiones de esta investigación podrían evaluar este potencial beneficio mediante un estudio riguroso utilizando métricas que evalúen el riesgo de las políticas de operación generadas.

Posibles mejoras para el modelamiento y las pruebas realizadas en esta investigación podrían enfocarse en la implementación de variantes de SDDP que aceleren la convergencia del algoritmo, que hagan un menor uso de memoria o reduzcan los tiempos de cálculo mediante una paralelización eficiente. Para mejorar el rendimiento del modelo basado en aprendizaje reforzado, se deben incorporar variables meteorológicas y utilizar arquitecturas con recursividades. En este trabajo no se llevó a cabo un estudio exhaustivo sobre arquitecturas neuronales, lo cual representa una oportunidad para mejorar el rendimiento del modelo basado en aprendizaje reforzado.

Adicionalmente, futuras investigaciones podrían explorar maneras de modelar con mayor precisión escenarios de precio. La representación markoviana de la incertidumbre abre la posibilidad de desarrollar un modelamiento de la incertidumbre más sofisticado, permitiendo incorporar correlaciones más complejas que las consideradas hasta ahora. Sin embargo, estas mejoras deben ir acompañadas de una mayor eficiencia en el algoritmo MC-SDDP. En relación a la toma de decisiones basadas en aprendizaje reforzado, resulta esencial estudiar la implementación de algoritmos de aprendizaje reforzado basados en modelo, que incluya un modelo detallado del BESS. Esta consideración es crucial para reducir los tiempos de cálculo y resolver el problema con una resolución de 5 minutos.

Bibliografía

- [1] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, “The role of renewable energy in the global energy transformation,” *Energy Strategy Reviews*, vol. 24, pp. 38–50, 2019.
- [2] C. Breyer, S. Khalili, D. Bogdanov, M. Ram, A. S. Oyewo, A. Aghahosseini, A. Gulagi, A. A. Solomon, D. Keiner, G. Lopez, P. A. Østergaard, H. Lund, B. V. Mathiesen, M. Z. Jacobson, M. Victoria, S. Teske, T. Pregger, V. Fthenakis, M. Raugei, H. Holtinen, U. Bardi, A. Hoekstra, and B. K. Sovacool, “On the history and future of 100 % renewable energy systems research,” *IEEE Access*, vol. 10, pp. 78176–78218, 2022.
- [3] IRENA and CPI, *Global landscape of renewable energy finance*. 2023.
- [4] M. Shafiullah, S. D. Ahmed, and F. A. Al-Sulaiman, “Grid integration challenges and solution strategies for solar pv systems: A review,” *IEEE Access*, vol. 10, pp. 52233–52257, 2022.
- [5] A. J. Dvorak and M. Victoria, “Key determinants of solar share in solar- and wind-driven grids,” *IEEE Journal of Photovoltaics*, vol. 13, no. 3, pp. 476–483, 2023.
- [6] O. M. Babatunde, J. L. Munda, and Y. Hamam, “A comprehensive state-of-the-art survey on hybrid renewable energy system operations and planning,” *IEEE Access*, vol. 8, pp. 75313–75346, 2020.
- [7] E. Pusceddu, B. Zakeri, and G. Castagneto Gisse, “Synergies between energy arbitrage and fast frequency response for battery energy storage systems,” *Applied Energy*, vol. 283, p. 116274, 2021.
- [8] G. He, Q. Chen, C. Kang, P. Pinson, and Q. Xia, “Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life,” *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2359–2367, 2016.
- [9] P. Barra, W. de Carvalho, T. Menezes, R. Fernandes, and D. Coury, “A review on wind power smoothing using high-power energy storage systems,” *Renewable and Sustainable Energy Reviews*, vol. 137, p. 110455, 2021.
- [10] E. Semshchikov, M. Negnevitsky, J. Hamilton, and X. Wang, “Cost-efficient strategy for high renewable energy penetration in isolated power systems,” *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3719–3728, 2020.

- [11] O. Palizban and K. Kauhaniemi, “Energy storage systems in modern grids—matrix of technologies and applications,” *Journal of Energy Storage*, vol. 6, pp. 248–259, 2016.
- [12] F. Fan, Y. Xu, R. Zhang, and T. Wan, “Whole-lifetime coordinated service strategy for battery energy storage system considering multi-stage battery aging characteristics,” *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 689–699, 2022.
- [13] W. Gil-González, O. D. Montoya, and J. C. Hernández, “An energy management system for the optimal operation of bess in dc microgrids: A robust convex programming approach,” *IEEE Access*, vol. 11, pp. 38168–38181, 2023.
- [14] A. Helseth and B. Mo, “Hydropower aggregation by spatial decomposition—an sddp approach,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 381–392, 2023.
- [15] A. Street, A. Brigatto, and D. M. Valladão, “Co-optimization of energy and ancillary services for hydrothermal operation planning under a general security criterion,” *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4914–4923, 2017.
- [16] R. B. S. Brandi, A. L. M. Marcato, B. H. Dias, T. P. Ramos, and I. C. da Silva Junior, “A convergence criterion for stochastic dual dynamic programming: Application to the long-term operation planning problem,” *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3678–3690, 2018.
- [17] P. Aaslid, M. Korpås, M. M. Belsnes, and O. B. Fosso, “Stochastic operation of energy constrained microgrids considering battery degradation,” *Electric Power Systems Research*, vol. 212, p. 108462, 2022.
- [18] M. V. F. Pereira and L. M. V. G. Pinto, “Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach,” *Water Resources Research*, vol. 21, no. 6, pp. 779–792, 1985.
- [19] M. V. F. Pereira and L. M. V. G. Pinto, “Multi-stage stochastic optimization applied to energy planning,” *Mathematical Programming*, vol. 52, pp. 359–375, May 1991.
- [20] M. Lehna, B. Hoppmann, C. Scholz, and R. Heinrich, “A reinforcement learning approach for the continuous electricity market of germany: Trading from the perspective of a wind park operator,” *Energy and AI*, vol. 8, p. 100139, 2022.
- [21] A. Fusco, D. Giofrè, A. Francesco Castelli, C. Bovo, and E. Martelli, “A multi-stage stochastic programming model for the unit commitment of conventional and virtual power plants bidding in the day-ahead and ancillary services markets,” *Applied Energy*, vol. 336, p. 120739, 2023.
- [22] H. Li, Z. Lu, Y. Qiao, B. Zhang, and Y. Lin, “The flexibility test system for studies of variable renewable energy resources,” *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1526–1536, 2021.

- [23] M. Zhang, N. Zhang, D. Guan, P. Ye, K. Song, X. Pan, H. Wang, and M. Cheng, “Optimal design and operation of regional multi-energy systems with high renewable penetration considering reliability constraints,” *IEEE Access*, vol. 8, pp. 205307–205315, 2020.
- [24] M. Al-Saffar and P. Musilek, “Reinforcement learning-based distributed bess management for mitigating overvoltage issues in systems with high pv penetration,” *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2980–2994, 2020.
- [25] P. González-Inostroza, C. Rahmann, R. Álvarez, J. Haas, W. Nowak, and C. Rehtanz, “The role of fast frequency response of energy storage systems and renewables for ensuring frequency stability in future low-inertia power systems,” *Sustainability*, vol. 13, no. 10, 2021.
- [26] S. F. Santos, M. Gough, D. Z. Fitiwi, A. F. P. Silva, M. Shafie-Khah, and J. P. S. Catalão, “Influence of battery energy storage systems on transmission grid operation with a significant share of variable renewable energy sources,” *IEEE Systems Journal*, vol. 16, no. 1, pp. 1508–1519, 2022.
- [27] M. Naemi, D. Davis, and M. J. Brear, “Optimisation and analysis of battery storage integrated into a wind power plant participating in a wholesale electricity market with energy and ancillary services,” *Journal of Cleaner Production*, vol. 373, p. 133909, 2022.
- [28] E. Bayborodina, M. Negnevitsky, E. Franklin, and A. Washusen, “Grid-scale battery energy storage operation in australian electricity spot and contingency reserve markets,” *Energies*, vol. 14, no. 23, 2021.
- [29] A. Attarha, P. Scott, and S. Thiébaux, “Network-aware co-optimisation of residential der in energy and fcas markets,” *Electric Power Systems Research*, vol. 189, p. 106730, 2020.
- [30] E. Chatterji and M. D. Bazilian, “Smart meter data to optimize combined roof-top solar and battery systems using a stochastic mixed integer programming model,” *IEEE Access*, vol. 8, pp. 133843–133853, 2020.
- [31] E. Achiluzzi, K. Kobikrishna, A. Sivabalan, C. Sabillon, and B. Venkatesh, “Optimal asset planning for prosumers considering energy storage and photovoltaic (pv) units: A stochastic approach,” *Energies*, vol. 13, no. 7, 2020.
- [32] L. Heistrene, B. Azzopardi, A. V. Sant, and P. Mishra, “Stochastic generation scheduling of insular grids with high penetration of photovoltaic and battery energy storage systems: South andaman island case study,” *Energies*, vol. 15, no. 7, 2022.
- [33] A. Soleimani, V. Vahidinasab, and J. Aghaei, “A linear stochastic formulation for distribution energy management systems considering lifetime extension of battery storage devices,” *IEEE Access*, vol. 10, pp. 44564–44576, 2022.
- [34] H. Park, “A stochastic planning model for battery energy storage systems coupled with utility-scale solar photovoltaics,” *Energies*, vol. 14, no. 5, 2021.

- [35] M. Grangereau, W. van Ackooij, and S. Gaubert, “Multi-stage stochastic alternating current optimal power flow with storage: Bounding the relaxation gap,” *Electric Power Systems Research*, vol. 206, p. 107774, 2022.
- [36] Z. Li, P. Yang, Y. Yang, G. Lu, and Y. Tang, “Hydropower preventive maintenance scheduling in a deregulated market: A multi-stage stochastic programming approach,” *Electric Power Systems Research*, vol. 221, p. 109379, 2023.
- [37] A. R. Silva, H. Pousinho, and A. Estanqueiro, “A multistage stochastic approach for the optimal bidding of variable renewable energy in the day-ahead, intraday and balancing markets,” *Energy*, vol. 258, p. 124856, 2022.
- [38] A. Shapiro, “Analysis of stochastic dual dynamic programming method,” *European Journal of Operational Research*, vol. 209, no. 1, pp. 63–72, 2011.
- [39] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2014.
- [40] A. Philpott and V. de Matos, “Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion,” *European Journal of Operational Research*, vol. 218, no. 2, pp. 470–483, 2012.
- [41] N. Löhdorf and A. Shapiro, “Modeling time-dependent randomness in stochastic dual dynamic programming,” *European Journal of Operational Research*, vol. 273, no. 2, pp. 650–661, 2019.
- [42] T. Asamov, D. F. Salas, and W. B. Powell, “Sddp vs. adp: The effect of dimensionality in multistage stochastic optimization for grid level energy storage,” 2016.
- [43] T. Asamov and W. B. Powell, “Regularized decomposition of high-dimensional multistage stochastic programs with markov uncertainty,” 2017.
- [44] F. D. Machado, A. L. Diniz, C. L. Borges, and L. C. Brandão, “Asynchronous parallel stochastic dual dynamic programming applied to hydrothermal generation planning,” *Electric Power Systems Research*, vol. 191, p. 106907, 2021.
- [45] D. Ávila, A. Papavasiliou, and N. Löhdorf, “Parallel and distributed computing for stochastic dual dynamic programming,” *Computational Management Science*, vol. 19, no. 2, pp. 199–226, 2022.
- [46] J. Zou, S. Ahmed, and X. A. Sun, “Stochastic dual dynamic integer programming,” *Mathematical Programming*, vol. 175, no. 1, pp. 461–502, 2019.
- [47] J. Zou, S. Ahmed, and X. A. Sun, “Multistage stochastic unit commitment using stochastic dual dynamic integer programming,” *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1814–1823, 2019.
- [48] S. Zhang and X. A. Sun, “Stochastic dual dynamic programming for multistage stochastic mixed-integer nonlinear optimization,” 2022.

- [49] D. Ávila, A. Papavasiliou, and N. Löhdorf, “Batch learning sddp for long-term hydrothermal planning,” *IEEE Transactions on Power Systems*, pp. 1–14, 2023.
- [50] F. Beltrán, E. Finardi, and W. de Oliveira, “Two-stage and multi-stage decompositions for the medium-term hydrothermal scheduling problem: A computational comparison of solution techniques,” *International Journal of Electrical Power & Energy Systems*, vol. 127, p. 106659, 2021.
- [51] A. Tilmant and R. Kelman, “A stochastic approach to analyze trade-offs and risks associated with large-scale water resources systems,” *Water Resources Research*, vol. 43, no. 6, 2007.
- [52] H. Macian-Sorribes, A. Tilmant, and M. Pulido-Velazquez, “Improving operating policies of large-scale surface-groundwater systems through stochastic programming,” *Water Resources Research*, vol. 53, no. 2, pp. 1407–1423, 2017.
- [53] F. Hafiz, A. Rodrigo de Queiroz, P. Fajri, and I. Husain, “Energy management and optimal storage sizing for a shared community: A multi-stage stochastic programming approach,” *Applied Energy*, vol. 236, pp. 42–54, 2019.
- [54] N. Löhdorf and D. Wozabal, “Gas storage valuation in incomplete markets,” *European Journal of Operational Research*, vol. 288, no. 1, pp. 318–330, 2021.
- [55] B. Yıldız and M. Sütçü, “A variant sddp approach for periodic-review approximately optimal pricing of a slow-moving a item in a duopoly under price protection with end-of-life return and retail fixed markdown policy,” *Expert Systems with Applications*, vol. 212, p. 118801, 2023.
- [56] V. Espanmanesh and A. Tilmant, “Optimizing the management of multireservoir systems under shifting flow regimes,” *Water Resources Research*, vol. 58, no. 6, p. e2021WR030582, 2022.
- [57] F. Cordera, R. Moreno, and F. Ordoñez, “Unit commitment problem with energy storage under correlated renewables uncertainty,” *Operations Research*, vol. 71, no. 6, pp. 1960–1977, 2023.
- [58] M. Anwar, C. Wang, F. de Nijs, and H. Wang, “Proximal policy optimization based reinforcement learning for joint bidding in energy and frequency regulation markets,” in *2022 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, 2022.
- [59] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *CoRR*, vol. abs/1707.06347, 2017.
- [60] H. Wang and B. Zhang, “Energy storage arbitrage in real-time markets via reinforcement learning,” in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, 2018.
- [61] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. A. Riedmiller, “Playing atari with deep reinforcement learning,” *CoRR*, vol. abs/1312.5602, 2013.

- [62] Y. Li, “Deep reinforcement learning,” *CoRR*, vol. abs/1810.06339, 2018.
- [63] A. Raffin, A. Hill, A. Gleave, A. Kanervisto, M. Ernestus, and N. Dormann, “Stable-baselines3: Reliable reinforcement learning implementations,” *Journal of Machine Learning Research*, vol. 22, no. 268, pp. 1–8, 2021.
- [64] H. Ding, Y. Xu, B. Chew Si Hao, Q. Li, and A. Lentzakis, “A safe reinforcement learning approach for multi-energy management of smart home,” *Electric Power Systems Research*, vol. 210, p. 108120, 2022.
- [65] J. Wei, X. Wu, T. Yang, and R. Jiao, “Ultra-short-term forecasting of wind power based on multi-task learning and lstm,” *International Journal of Electrical Power & Energy Systems*, vol. 149, p. 109073, 2023.

ANEXOS

Anexo A

Mercado eléctrico australiano

El NEM es un mercado que suministra energía eléctrica a 5 estados: Queensland (QLD), New South Wales (NSW), Victoria (VIC), South Australia (SA) y Tasmania (TAS). El mercado opera con una resolución de 5 minutos y en él se oferta energía y servicios auxiliares de control de frecuencia. Con las ofertas recibidas el día anterior, se co-optimiza el suministro eléctrico y los servicios auxiliares, con el objetivo operar a mínimo costo y satisfacer un forecast de demanda con seguridad N-1 [27].

En cuanto a los servicios auxiliares de control de frecuencia, por el momento existen 8: 2 de regulación, destinados a corregir desbalances menores entre la generación y la demanda; y 6 de contingencias, destinados a evitar desviaciones mayores en la frecuencia en caso de contingencias en la red. Los servicios de regulación están compuesto por uno de bajada y otro de subida. Mientras que los servicios ante contingencias pueden ser de bajada o subida, y operar en ventanas de 6 segundos, 60 segundos o 5 minutos [27]. En la tabla A.1, recuperada de [28], se presenta un resumen de los servicios auxiliares de control de frecuencia que forman parte del NEM y las condiciones en las que opera cada uno .

Tabla A.1: Mercados australianos de servicios auxiliares de control de frecuencia.

Mercado FCAS		Función	Descripción
Regulación	Bajada	Corregir alza/baja menor en frecuencia.	Orientado a responder a desviaciones menores a $\pm 0,15$ Hz.
	Subida		
Contingencia	Bajada rápida	Respuesta en una ventana de 6s. La respuesta debe sostener por 60s.	Orientado a responder a contingencias cuando la frecuencia se desvía del NOB.
	Subida rápida		
	Baja lenta	Respuesta en una ventana de 60s. La respuesta debe sostener por 300s.	
	Subida lenta		
Bajada con retraso	Respuesta en una ventana de 5min. La respuesta debe sostenerse por 10min.		
Subida con retraso			

Anexo B

Resultados

B.1. Arbitraje de energía mediante optimización

La figura B.1 presenta las utilidades obtenidas al variar la resolución con la que opera el mercado, relativas al modelo markoviano con resolución de 5 minutos y sin desestacionalizar precios. Se observa que la brecha en las utilidades del modelo markoviano y el determinista se reduce cuando los precios no se desestacionalizan.

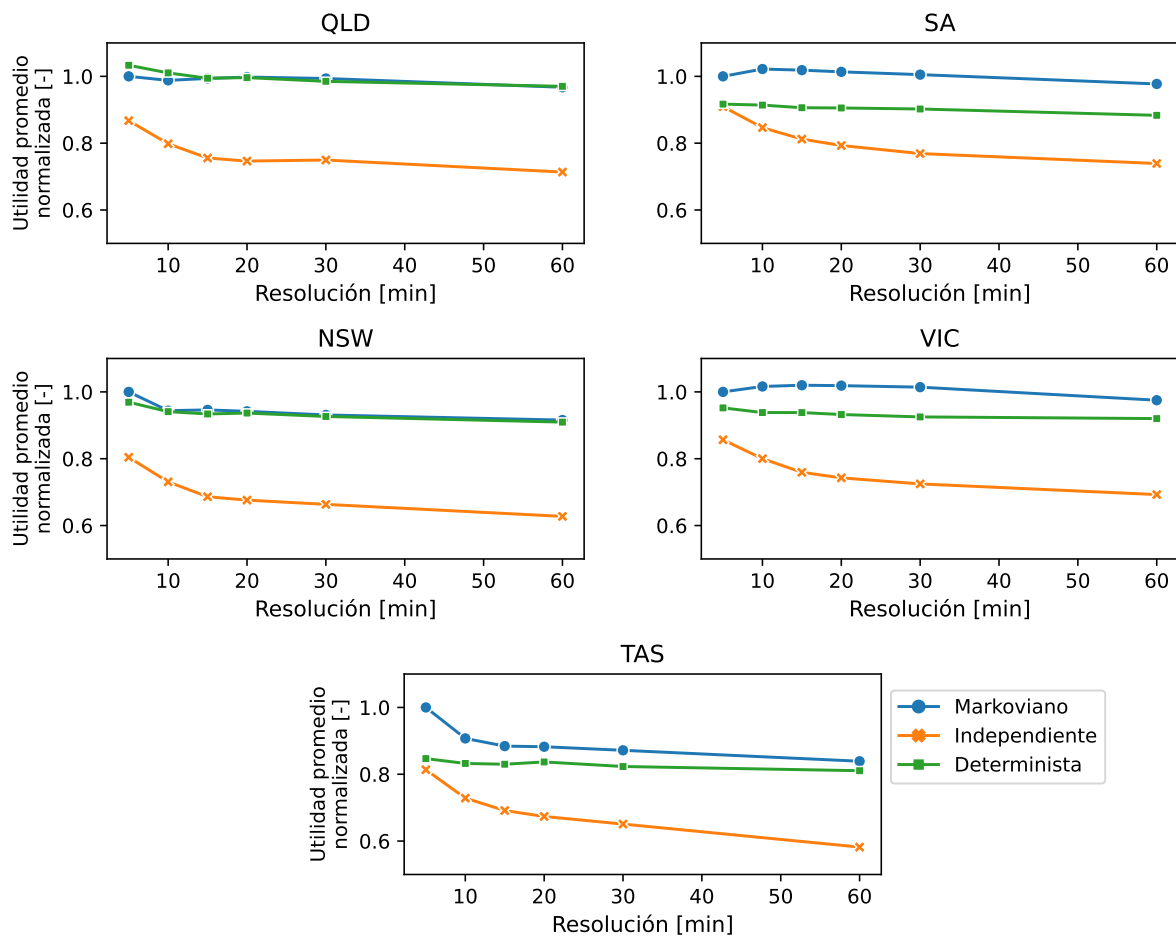


Figura B.1: Utilidades relativas de los modelos de optimización al arbitrar energía. No se desestacionalizan los precios y se consideran diferentes resoluciones en la operación del mercado.

La figura B.2 muestra las utilidades promedio para los modelos basados en optimización al arbitrar energía, tanto si se desestacionaliza como si no se desestacionalizan los precios de la energía. Además, se presentan las utilidades de un modelo que revela la incertidumbre en la primera etapa. Aunque este modelo carece de realismo, su utilidad reside en confirmar que las utilidades son iguales tanto si no se desestacionalizan los precios de la energía como si se hace y luego se reestacionalizan las utilidades.

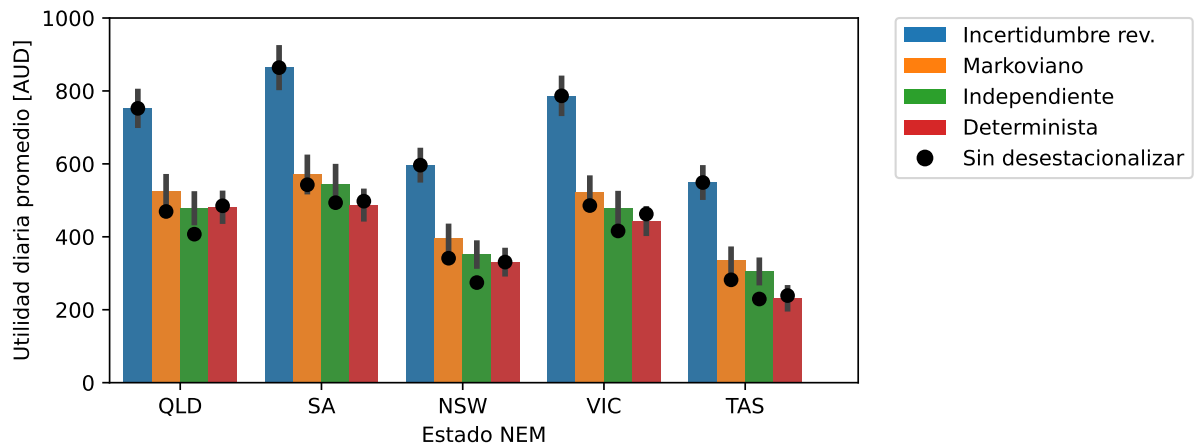


Figura B.2: Utilidades de los modelos basados en optimización al arbitrar energía, mercado con resolución de 5 minutos. Las barras representan las utilidades corregidas al usar precios desestacionalizados. Los puntos representan las utilidades al no desestacionalizar precios.

B.2. Arbitraje de energía y participación en el mercado de reservas mediante optimización

En la figura B.3 se presenta el precio promedio de las reservas junto con el rango interdecil (rango entre el decil 1 y el decil 9) del año 2022. Se observa que los mercados de 5 minutos presentan los precios más bajos.

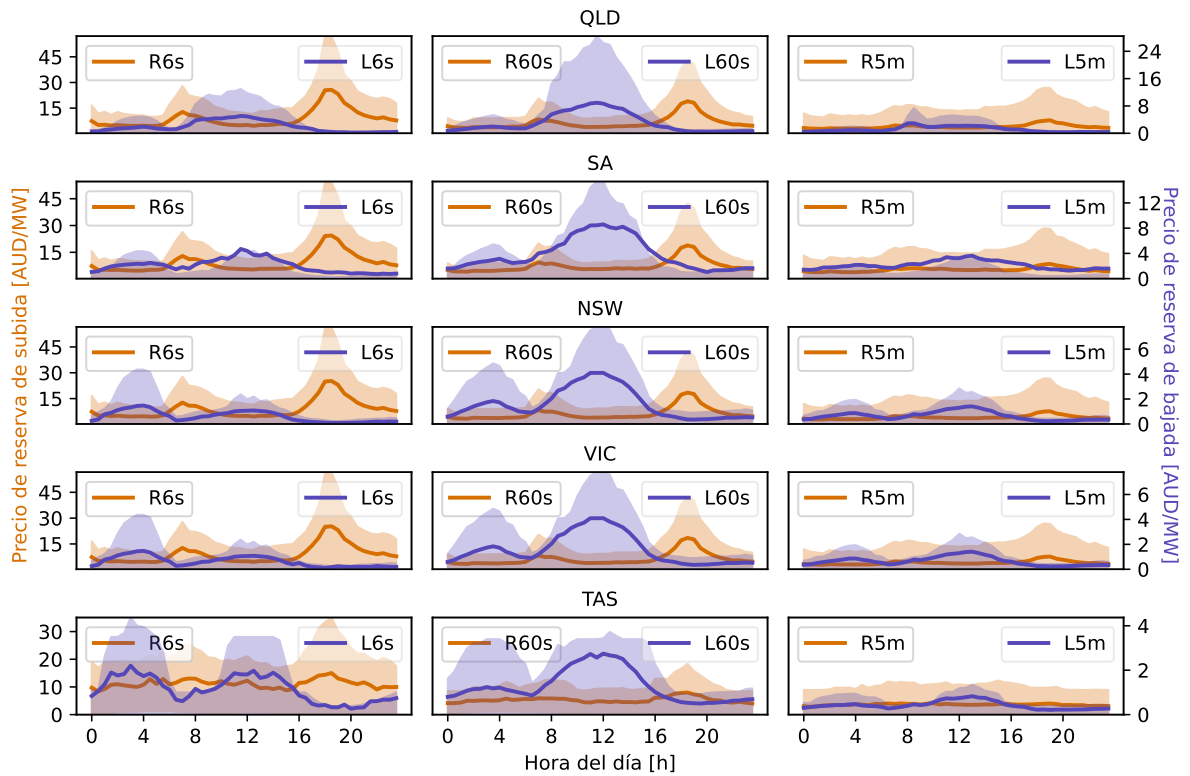


Figura B.3: Precio promedio y rango interdecil horario de los precios de las reservas del año 2022. Se han excluido aquellas curvas de precio que presentan valores superiores a 600[AUD/MW] o inferiores a -600 [AUD/MW].

En la figura 4.6 se presenta las utilidades relativas obtenidas al usar el modelo propuesto en la metodología. A diferencia del cuerpo del informe, para elaborar esta figura se excluyen los días con utilidades superiores a 10.000 [AUD].

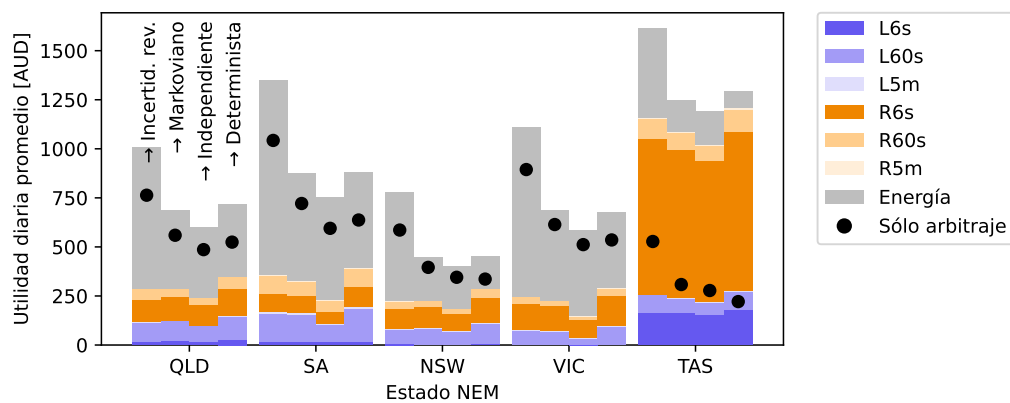


Figura B.4: Utilidades de los modelos de optimización al participar en el mercado de reservas y arbitrar energía, mercado con resolución de 10 minutos y precios desestacionarizados. Se excluyen los días con utilidades superiores a 10.000[AUD].

B.3. Arbitraje de energía mediante aprendizaje reforzado

La figura B.5 presenta las utilidades obtenidas al utilizar modelos basados en optimización y en aprendizaje reforzado.

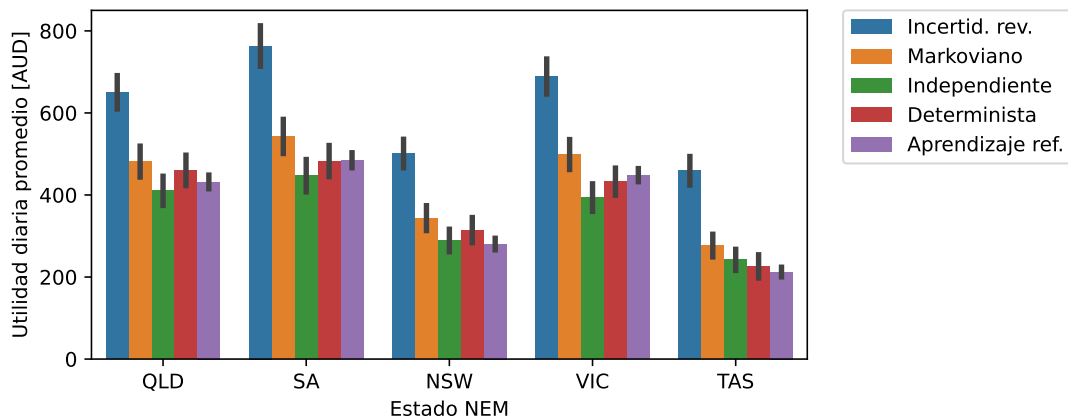


Figura B.5: Utilidades de los modelos de optimización y el basado en aprendizaje reforzado al realizar arbitraje de energía, mercado con resolución de 30 minutos. Las utilidades de los modelos basados en optimización se reestacionalizan.

La Tabla B.1 proporciona información sobre el tiempo de cálculo y el uso de memoria de los modelos de toma de decisiones empleados en este estudio. Los cálculos se realizaron en un equipo con 64 GB de memoria RAM y un procesador AMD Ryzen 7 5800X3D. En el contexto de arbitraje energía en resolución de 30 minutos, es notable que los modelos basados en aprendizaje reforzado consumen una cantidad de recursos considerablemente mayor que cuando el problema es resuelto empleando MC-SDDP. Específicamente, el uso de memoria aumenta en 5 veces y el tiempo de cálculo en cerca de 100. Esto ocurre pues el modelo basado en aprendizaje reforzado no cuenta con un modelo matemático del BESS, requiriendo un elevado número de iteraciones para capturar las relaciones entre inyección de potencia, estado de carga e ingresos.

Tabla B.1: Tiempo de computo y uso de memoria de los modelos de toma de decisión tratados en esta investigación. Los valores indicados representan los recursos empleados durante una única ejecución de entrenamiento, y no deben interpretarse como valores promedio.

Problema	Algoritmo	Resolución [min]	Tiempo de entrenamiento [min]	Uso máximo de memoria [GB]
Arbitraje	MC-SDDP	60	0,3	0,4
		30	0,7	0,5
		20	1,0	0,7
		15	1,5	0,9
		10	2,2	1,3
		5	4,1	2,5
	SDDP	60	0,2	0,15
		30	0,3	0,17
		20	0,5	0,17
		15	0,6	0,20
		10	1,0	0,23
		5	1,9	0,25
	Determinista	60	< 0,1	0,13
		30	< 0,1	0,14
		20	< 0,1	0,13
		15	< 0,1	0,13
		10	< 0,1	0,13
		5	< 0,1	0,15
RL	30	51	2,5	
Arbitraje y reservas	MC-SDDP	10	37	13
		5	78	23
	SDDP	10	13	1,5
		5	46	4,7
	Determinista	10	< 0,1	0,14
		5	< 0,1	0,16

Las figuras B.6 y B.7 muestran la operación del BESS al utilizar modelos basados en optimización y aprendizaje reforzado, respectivamente. Los datos empleados corresponden a las primeras dos semanas del 2022 del bus de referencia ubicado en Tasmania, y la operación se realiza para un horizonte de 24 horas (cada 24 horas la batería inicia con el mínimo estado de carga posible). Al igual que con la operación en el nodo de referencia en Queensland, el modelo determinista y el markoviano presentan dos ciclos de carga-descarga; sin embargo, en este caso los ciclos son más irregulares.

Respecto al modelo basado en aprendizaje reforzado, la mayoría de las veces se observan dos ciclo de carga-descarga y en los mismos horarios que al utilizar modelos de optimización. Esta situación podría estar influenciada por la forma de la curva de demanda. Adicionalmente, en la mayoría de los días de operación, el estado de carga no llega a su máximo permitido, lo que podría indicar que no se está aprovechando la totalidad de

la capacidad de almacenamiento.

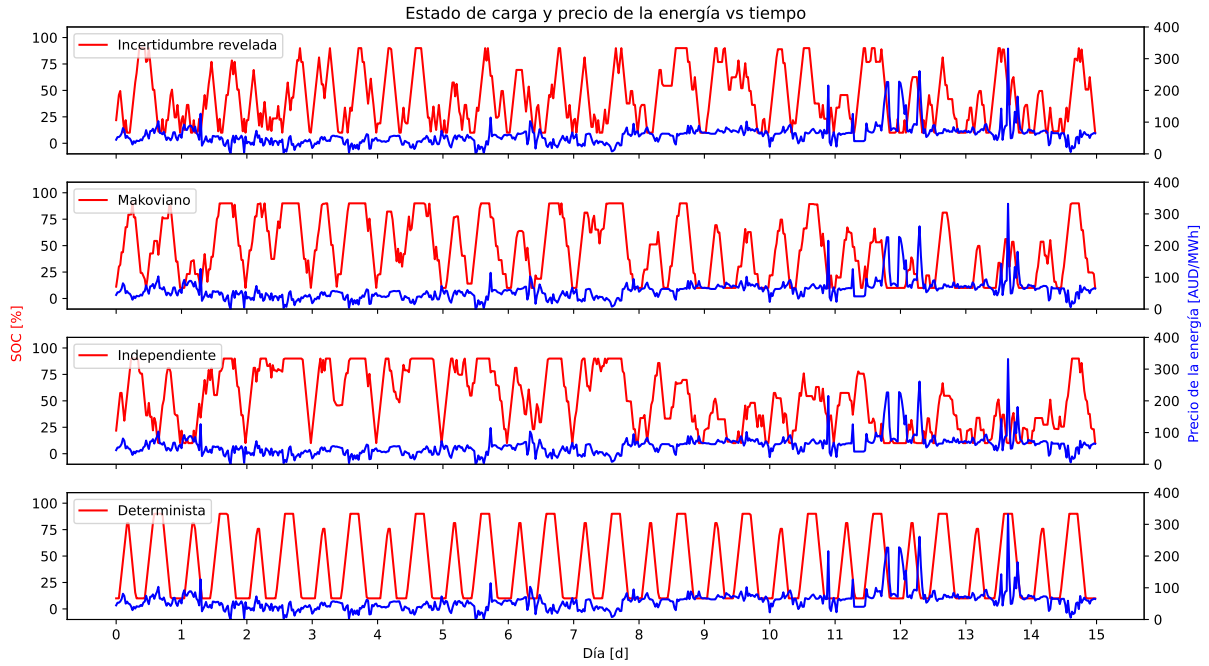


Figura B.6: Operación de los primeros 15 días del año 2022 al utilizar modelos basados en optimización para arbitrar energía. Datos del nodo de referencia ubicado en Tasmania.

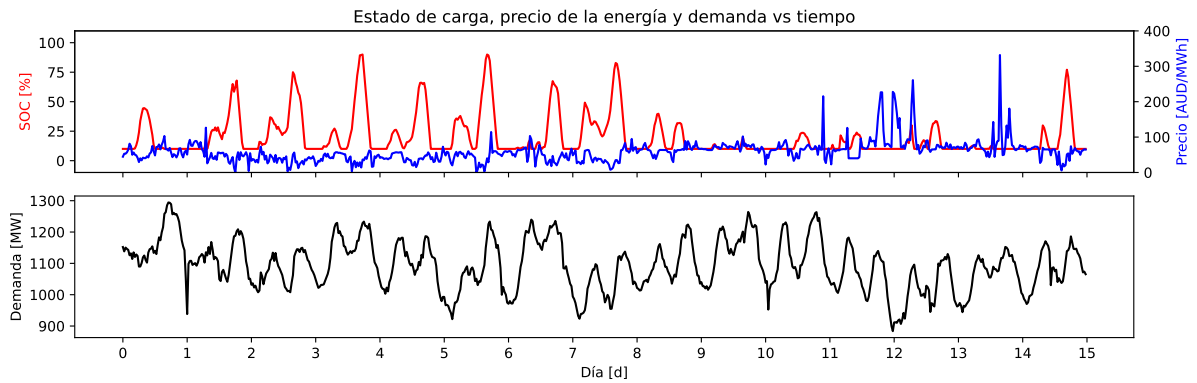


Figura B.7: Operación de los primeros 15 días del año 2022 al utilizar aprendizaje reforzado para arbitrar energía. Datos del nodo de referencia ubicado en Tasmania.