



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

DIMENSIONAMIENTO ÓPTIMO DE SISTEMAS DE ALMACENAMIENTO BASADO EN
TECNICAS DE REDUCCION DE ESCENARIOS.

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

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

MIGUEL ALEXIS SÁNCHEZ LÓPEZ

PROFESOR GUÍA:
RODRIGO ERNESTO PALMA BEHNKE

MIEMBROS DE LA COMISIÓN:
PABLO ANTONIO ESTEVEZ VALENCIA
ESTEBAN GIL SAGAS

Este trabajo ha sido financiado por Conicyt

SANTIAGO DE CHILE
2018

RESUMEN DE LA TESIS PARA OPTAR AL TÍTULO DE: GRADO
DE MAGÍSTER EN CIENCIAS DE LA ING. MENCIÓN
ELÉCTRICA Y AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: MIGUEL ALEXIS SÁNCHEZ LÓPEZ
FECHA: 12/01/2018
PROFESOR GUÍA: RODRIGO PALMA BEHNKE.

DIMENSIONAMIENTO ÓPTIMO DE SISTEMAS DE ALMACENAMIENTO BASADO EN TÉCNICAS DE REDUCCIÓN DE ESCENARIOS

El presente trabajo presenta una novedosa metodología que permite realizar reducción de escenarios en problemas de planificación de la generación incluyendo almacenamiento de energía en sistemas eléctricos de potencia. La metodología se justifica por el hecho de que la planificación de sistemas eléctricos con alta penetración de fuentes de energía variable incita al planificador a tener en cuenta las restricciones de corto plazo, las que son componente habitual en los modelos de predespacho de unidades “*unit commitment*”. Adicionalmente, la variabilidad de las fuentes renovables ha impulsado la incorporación de sistemas de almacenamiento como una manera de aportar a la flexibilidad requerida para conseguir el nivel de confiabilidad y seguridad necesario para operar el sistema.

Por otra parte, la naturaleza incierta que poseen las fuentes variables hace natural que el problema de optimización subyacente sea de carácter estocástico, el cual se enfrenta mediante escenarios en un árbol determinístico equivalente. Como consecuencia, se tiene un modelo cuya resolución requiere grandes esfuerzos computacionales. Considerando los elementos mencionados, esta tesis propone un manejo eficiente de la incertidumbre mediante la reducción de la cantidad de escenarios, asegurando a su vez un nivel predeterminado de precisión y exactitud en la solución encontrada.

La metodología utilizada propone cuatro fases. Las dos primeras denominadas de Selección de Características y Reducción de Escenarios, se basan en métricas de teoría de la información. El propósito de ambas fases es caracterizar cada escenario según la manera en que estos afectan a la planificación. Con ese objetivo se evalúan características de la salida del problema de optimización que modela la planificación, y de este conjunto se escoge aquella que mejor represente la variable aleatoria. Posteriormente, el conjunto de escenarios se reduce aplicando una técnica de *clustering* cuya función objetivo está enfocada en que las distribuciones de probabilidad de la cantidad original de escenarios y la versión reducida sean cercanos en un espacio de probabilidad. Una tercera fase de Análisis de Post-optimalidad se encarga de analizar la bondad de los resultados obtenidos en la planificación con los escenarios reducidos, y dependiendo de las conclusiones obtenidas, se procede a realizar ajustes de la técnica en una cuarta fase, llamada ajustes.

Con el objetivo de validar el funcionamiento de cada fase y el potencial de la metodología se analizan dos casos de estudio, los que corresponden a una planificación de la generación y almacenamiento con un horizonte de una semana y un semestre. Este último da cuenta de las proyecciones del sistema eléctrico chileno al año 2030. Los resultados obtenidos muestran que la metodología, puede llegar a valores cuyo error estimado es menor a un 2% respecto a los costos totales del sistema utilizado un 3% de los datos. Más aún, se proponen estimadores para determinar la inversión en tecnologías de almacenamiento. Finalmente, dicha metodología es capaz de realizar una estimación en la inversión incluso en contextos donde ejecutar un modelo con la totalidad de escenarios es computacionalmente no factible. Como trabajo futuro se propone explotar el potencial de paralelización e incorporar el sistema de transmisión a los modelos de planificación. Asimismo, para cada fase del método propuesto se identifican mejoras específicas.

Dedicatoria

A mis padres Carmen y Miguel.

Agradecimientos.

Trabajar en esta tesis significó trabajar para cerrar una etapa en mi vida. Etapa que partió y culminó con mis padres Carmen y Miguel dando todo para que mi única preocupación en la vida sea cumplir mis sueños. Me han hecho la vida fácil y todo lo que pueda decir nunca será suficiente para demostrar mis sentimientos. El mérito de mis éxitos es todo de ustedes y solo de ustedes. No quiero olvidar a mis hermanas Bárbara y Karin, por recordarme quien soy en aquellos momentos en que sentía que las cosas se complicaban, mi cariño es infinito.

Nada de lo escrito en estas páginas sería posible sin la enorme ayuda del Profesor Rodrigo Palma. Solo redundaría si me refiriera a sus enormes capacidades docentes, pero eso está lejos de ser lo que más valoro. También sería cliché decir que agradezco su confianza depositada en mí durante estos años, aunque por cierto, lo hago. Sin embargo, lo que jamás olvidaré es su capacidad por hacer que yo confíe en mí mismo como nunca antes lo había hecho, mi enorme gratitud y respeto por todo ello. También agradecer el apoyo a los profesores Pablo Estévez, Marcelo Cortes y Esteban Gil por sus recomendaciones y disponibilidad para apoyarme en esta tesis.

A los que trabajaron directamente conmigo en este tema, Sebastián Morris, quien su cariño y amistad es de las cosas que más valoro de aprender a reducir escenarios. Y por supuesto, a Francisco Celis, quien fue un gran compañero para discutir acerca del almacenamiento, pero mucho más que eso, es un gran amigo que me deja este proceso, gracias porque siento que he aprendido mucho de tu calidad humana y nobleza.

Pero, ¿Cómo terminar los agradecimientos sin abordar a aquellos que me han hecho crecer como persona? Y es que con aquellos aprendí que la gran ventaja de un paso rápido a castillos no está en el doble establo, si no en las amistades que se forjan. O entender de que el mundo es nuestro y que para ello no existen fronteras, porque eso fue siempre cierto, cierto en la ebria, cierto en el Esperanto, en el Diablito, cierto tomando una TsingTao caliente en la madrugada de Xi'an, cierto en la Ovo, cierto en Time Square y cierto en cada lugar que esté con ustedes. De quien aprendí la sencillez e importancia de ser feliz todos los días y que bailando todo se arregla, y vaya que todo es más hermoso. A quien conoció mis peores caras, pero aun así solo bastó una junta, un partido de squash un viernes a las 9 para darme cuenta de lo enorme que es y lo afortunado que soy de tenerte como amiga. A los incondicionales que estuvieron ahí para darme el consejo oportuno, comiendo un as, yendo al cine o pedaleando un sábado en la mañana. Y los que estuvieron siempre conmigo, para decirme lo que debía oír y no lo que quería escuchar, a los creadores del algoritmo, a los compañeros de mil batallas, a mis amigos condor, merlot y capo del trap, porque con ustedes el partido jamás lo jugué solo.

Finalmente, quiero agradecer al proyecto Fondecyt 1151438 por financiar esta investigación.

Tabla de contenido

1. Introducción	1
1.1. Hipótesis de la propuesta	4
1.2. Objetivos y alcances.....	4
1.3. Resultados esperados	5
1.4. Estructura del documento.....	5
2. Marco Teórico.....	6
2.1. Planificación de sistemas eléctricos con almacenamiento: Problemáticas y desafíos	6
2.1.1. Notación de la sección	7
2.1.2. Función objetivo	9
2.1.3. Restricciones del problema de planificación considerando limitantes de corto plazo.....	10
2.2. Métricas de información a partir de muestras	14
2.2.1. Estimación de densidades de probabilidad mediante ventanas de Parzen	14
2.2.2. Entropía, divergencia e información mutua.....	15
2.3. Programación estocástica, análisis de post optimalidad	17
2.3.1. Programación estocástica	18
2.3.2. Técnica de los contaminantes	20
2.4. Algoritmos de reducción de escenarios	22
3. Metodología	25
3.1. Nomenclatura del capítulo	25
3.2. Fase I: Selección de características	28
3.3. Fase II: Reducción de escenarios	28
3.4. Fase III: Análisis de post optimalidad.....	30
3.5. Fase IV: Ajustes	32
3.6. Comparación con algoritmos en la literatura	33
4. Casos de estudio.....	35
4.1. PGA de prueba	35

4.2. Modelo PGA aplicado al Sistema Eléctrico Nacional chileno	37
5. Resultados y Análisis	41
5.1. PGA de prueba	41
5.1.1. Fase I: Selección de características	43
5.1.2. Fase II: Reducción de escenarios	45
5.1.3. Fase III: Análisis de post optimalidad y Fase IV: Ajustes	48
5.1.4. Revisión general del PGA de prueba.	50
5.2. PGA aplicado al SEN.....	50
5.2.1. Aplicación de la metodología en el modelo del SEN chileno.....	51
6. Conclusiones y Trabajo Futuro.....	55
6.1. Conclusiones sobre la metodología.....	55
6.2. Trabajo Futuro.....	56
7. Bibliografía	59

Índice de tablas

Tabla 3-1 Comparación de la propuesta metodológica con los algoritmos de la literatura	34
Tabla 4-1 Parámetros de tecnologías de generación modelo PGA de prueba.	35
Tabla 4-2 Parámetros de los sistemas de almacenamiento modelo PGA de prueba.....	36
Tabla 4-3 Fuentes de los datos modelo PGA aplicado al SIN.	37
Tabla 4-4 Parámetros técnico económicos centrales térmicas del modelo PGA aplicado al SIN.	38
Tabla 4-5 Parámetros técnicos económicos sistemas de almacenamiento modelo PGA aplicado al SIN.	38
Tabla 4-6 Parámetros técnico económicos de tecnologías de FEV modelo PGA aplicado al SIN	39
Tabla 5-1 Información mutua entre las series de viento y características de PGA de prueba.	43
Tabla 5-2 Índices de adaptabilidad para el PGA de prueba.	49
Tabla 5-3 Comparación solución de reducción de escenarios y cantidad total de escenarios.	50
Tabla 5-4 Información mutua entre las series de viento y características del SEN chileno.	51
Tabla 5-5 Índices de adaptabilidad para la reducción de escenarios a $m = 3$	53
Tabla 5-6 Índices de adaptabilidad para la reducción de escenarios a $m = 4$	54

Índice de ilustraciones

Ilustración 1-1 Diferencia entre soluciones precisas y soluciones exactas	2
Ilustración 1-2 Esquema de Programación de la Generación y Almacenamiento estocástico	3
Ilustración 2-1 Representación gráfica de un problema estocástico.	20
Ilustración 3-1 Diagrama de la propuesta metodológica.	27
Ilustración 4-1 Perfil demanda y solar modelo PGA de prueba.	36
Ilustración 4-2 Ejemplos de serie de viento modelo PGA de prueba.	36
Ilustración 4-3 Perfil de demanda y disponibilidad solar de una semana modelo PGA aplicada al SIN.	39
Ilustración 4-4 Disponibilidad Eólica modelo PGA aplicado al SIN.	40
Ilustración 5-1 Mix promedio sistema de PGA determinísticos de prueba.	41
Ilustración 5-2 Desviación estándar asociada al mix óptimo aplicado al PGA de prueba.....	42
Ilustración 5-3 Mix óptimo del PGA estocástico aplicado al caso de prueba.....	43
Ilustración 5-4 Densidad de probabilidad de las características normalizadas (PGA de prueba)..	45
Ilustración 5-5 Divergencia de Cauchy Schwarz para el caso de reducción a 3 escenarios (PGA de prueba)	46
Ilustración 5-6 Convergencia del gradiente descendente para el caso PGA de prueba.	47
Ilustración 5-7 Comparación de densidad de probabilidad y su reducción óptima con la encontrada con el algoritmo.	48
Ilustración 5-8 Análisis de post optimalidad PGA de prueba.	49
Ilustración 5-9 Mix óptimo obtenido con 3 escenarios aplicado al SEN chileno.	52
Ilustración 5-10 Despacho económico de dos semanas del mix óptimo alcanzado.....	52

1. Introducción

Durante los últimos años, diversos hechos como son el crecimiento de la demanda energética a nivel mundial, el calentamiento global y el desarrollo de políticas sustentables han cultivado el interés por el desarrollo y estudio del uso de energías renovables. Sin ir más lejos, durante el 2016 la capacidad instalada en generación renovable tuvo su mayor incremento en la historia, con 161 GW incorporados a nivel mundial. Dicha cifra corresponde al 62% del total de la capacidad que se instaló en dicho año [1].

Sin embargo, el uso intensivo de estas tecnologías trae una serie de desafíos relacionados con la operación de un sistema eléctrico de potencia, hecho que se acentúa cuando el recurso renovable proviene de fuentes de energía variable (FEV) y con incertidumbre en su disponibilidad, como lo son la energía eólica y la energía solar [2]. Cabe señalar que la característica *variabilidad* está relacionada con cambios abruptos en la disponibilidad del recurso, pudiendo ésta cambiar en una pequeña ventana de tiempo. La *Incertidumbre*, por su parte, está relacionado con los errores que se pueden cometer en el pronóstico que se tenga sobre la disponibilidad de la FEV, siendo más inciertas aquellas que se espera tener un mayor error asociado.

Es entonces que ante una mayor participación de FEV en un sistema eléctrico de potencia, éste incita a que la red tenga la capacidad de operar con la flexibilidad y resiliencia necesaria de modo de mantener niveles adecuados de confiabilidad. Por tanto, en ese contexto es que los sistemas de almacenamiento emergen como una de las principales alternativas que contribuyen el desarrollo de sistemas eléctricos con alta penetración de FEV [3]. Entre distintos modelamientos que se exhiben en la literatura, se encuentra el de la referencia [4], en el cual se busca maximizar la utilidad de una batería a nivel de distribución, la cual ofrece tres tipos de servicio relacionados con servicios de energía (realizando arbitraje de precios), transmisión (al realizar *Peak Shaving*), y flexibilidad (pudiendo participar en el mercado de reservas). Un mayor detalle del estado del arte en cuanto al uso de sistemas de almacenamiento se presenta en [5].

Por otro lado, la planificación de un sistema eléctrico esencialmente térmico, hasta hace unas décadas atrás, se realizaba desde un enfoque en que la demanda era representada mediante una curva de duración de carga, perdiendo la cronología e ignorando fenómenos del corto plazo como son los costos de encendido y apagado de las unidades, mínimos tiempos de operación y apagado, tasas de toma de carga [6]. Sin embargo, en sistemas con alta penetración de FEV, ignorar las restricciones de corto plazo puede implicar llegar a soluciones sub-óptimas, debido al incremento de los costos operacionales que se debe enfrentar en tiempo real [7]. Es más, en la literatura se han reportado diferencias entre el 7% y 17% en los costos de operación para el caso del SING chileno, y el sistema eléctrico de Texas respectivamente. Los sistemas de almacenamiento, dependiendo de la complejidad del modelo son representados a través de ecuaciones generales de balance energético. Sin embargo, restricciones adicionales pueden estar asociadas a la capacidad de rampa de la tecnología, o a mínimos estados de carga para considerar el desmedro del estado de salud de una batería [8].

Más aun, la naturaleza estocástica de las FEV impone la necesidad de tener que lidiar de algún modo con la incertidumbre en los modelos de Planificación de la Generación y Almacenamiento (PGA). Se distinguen muchas variantes para realizar una PGA mediante programación estocástica, muchas de las cuales se presentan en las referencias en [9]–[11]. La ventaja de este tipo de modelación es la incorporación endógena de la incertidumbre y variabilidad del recurso, a diferencia de modelos determinísticos, los cuales tratan de incorporar la incertidumbre mediante reglas que se traducen en restricciones con parámetros exógenos. Del mismo modo, el estudio

realizado en [12] muestra el impacto de planificar un sistema eléctrico con distintas funciones objetivo que expresan distintos grados de aversión al riesgo.

Sin embargo, una PGA estocástica con restricciones de corto plazo requiere una carga computacional importante. Entre las posibles alternativas para mitigar dicha carga está el uso de técnicas de descomposición como es la descomposición de Benders [13] [14]. Sin embargo, cualquiera sea la técnica de descomposición, la carga computacional dependerá fuertemente de la cantidad de escenarios que se usen.

En este punto, es conveniente notar la diferencia entre soluciones precisas y soluciones exactas (Ilustración 1-1). Soluciones *precisas* de un problema estocástico son aquellas que tienen poca varianza, sin embargo, pueden tener un sesgo. Soluciones *exactas*, por su parte, son aquellas que la esperanza de su error será pequeña o cero, pero que, sin embargo, podrían tener una gran varianza. De modo que idealmente, las soluciones de un problema estocástico deberían responder a ambas nociones

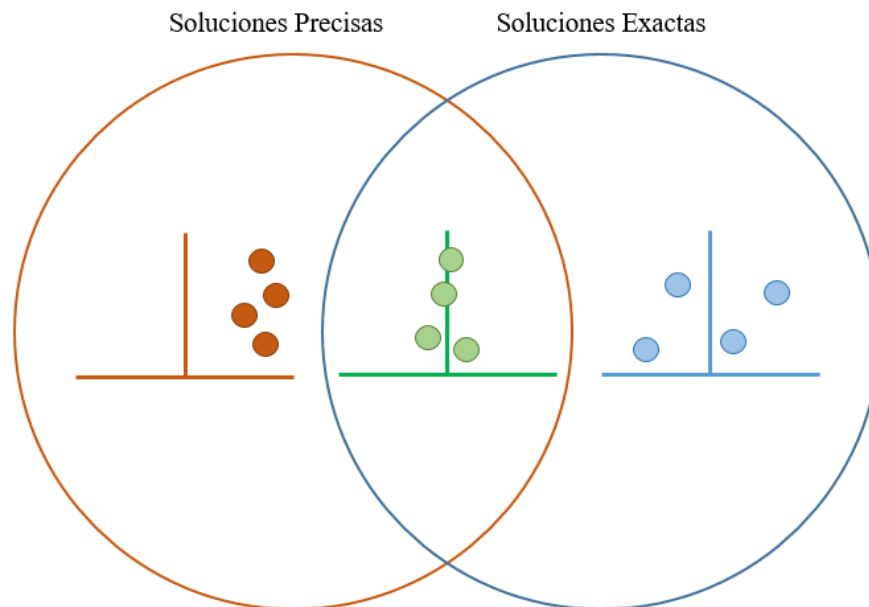


Ilustración 1-1 Diferencia entre soluciones precisas y soluciones exactas

La Ilustración 1-2 describe como los escenarios son representados en una estructura de árbol, donde cada hoja representa un escenario, y a su vez tienen un único camino al nodo raíz. Cada nodo representa una etapa de decisión, la cual depende de las decisiones anteriores y del escenario del nodo en particular (lo que se conoce como el principio de no anticipación). Cada escenario representa una realización de una variable aleatoria como puede ser la disponibilidad de las unidades en el sistema eléctrico de potencia, el precio de los combustibles, o la disponibilidad de los FEV etc. En general, los escenarios consisten en un conjunto numerable $\{\omega_i\}_{i \in I}$ de n realizaciones de una variable aleatoria continua P . Si n es suficientemente grande, la función distribución de masas P_ω converge a P , y la solución de ambos PGA con respecto a ambas distribuciones tiende a ser igual [15]. Más aun, bajo ciertos supuestos, un problema estocástico multi-etapas $F(P_\omega)$ es una función Lipschitz continua, lo que significa que la distancia entre las soluciones de $F(P_\omega)$ y $F(Q_\omega)$ son proporcionales a la distancia entre P_ω y Q_ω en el espacio de

probabilidades. Es entonces que dos distribuciones cercanas responderán a la noción de exactitud, en el sentido de que el valor de su función objetivo se moverá en una vecindad en torno a la solución.

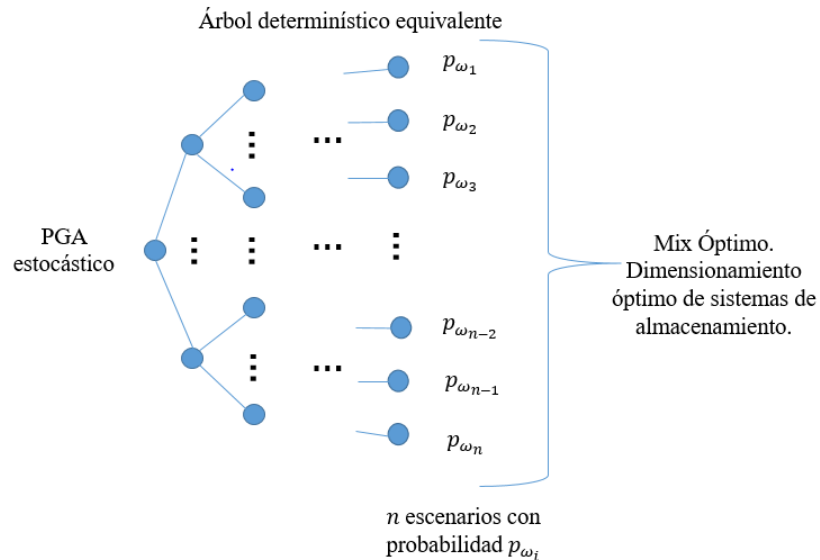


Ilustración 1-2 Esquema de Programación de la Generación y Almacenamiento estocástico

Así, para conseguir soluciones exactas, basta un número suficientemente grande de realizaciones de la variable aleatoria en cuestión. Sin embargo, en muchas ocasiones no se cuenta con la capacidad de cómputo necesaria para abordar un problema de esas dimensiones. Entonces es posible seleccionar un subconjunto de m muestras del conjunto original $\{\omega_i\}_{i \in I}$ y sus probabilidades asociadas. Este enfoque recibe el nombre de Reducción de Escenarios.

En este contexto, se reconocen dos tipos de reducción de escenarios. La gran mayoría de publicaciones centra sus esfuerzos en reducir escenarios con el objetivo de minimizar la distancia de los escenarios originales y los escenarios reducidos, pero desde una perspectiva del espacio de probabilidades sin consideración del contexto que impone el problema de optimización. La referencia [16] sienta las bases para este tipo de reducción de escenarios y el resto de las referencias suele citarla. El estudio [17] una técnica basada en minimizar la distancia de Kantorovic es propuesta con el objetivo de reducir escenarios para un problema de optimización multi-etapa. Finalmente, trabajos recientes han estudiado los límites de validez y garantías teóricas de este tipo de reducción de escenario [18]. Por otra parte [19] propone un método de reducción de escenarios en que evalúa la bondad de los resultados desde el punto de vista del valor de la función objetivo y también de las variables de decisión.

El problema de reducción óptima de escenarios en los métodos recién mencionados es un problema *NP-hard*. Por dicho motivo, distintos algoritmos heurísticos han sido desarrollados para tratar con este problema, por ejemplo, el *forward selection* y *backward reduction* [16]. En la referencia [20] se presenta una optimización por Enjambre de Partículas para reducir escenarios buscando minimizar un criterio de distancia.

Por otro lado, una menor cantidad de publicaciones se han enfocado en el desarrollo de técnicas basadas en caracterizar cada escenario según la manera que esto afecta a la salida del problema de planificación de la expansión (por ejemplo, la capacidad instalada en alguna tecnología, costos totales, etc.). La referencia [21] propone una reducción de escenarios basado en una técnica de

clustering, la cual caracteriza cada escenario con los costos totales de la política de planificación calculados de una manera determinística.

Incluso cuando la distribución de los m escenarios seleccionados a partir de n escenarios originales ($m < n$) tenga un valor óptimo considerando alguna métrica particular, la pregunta de qué tan precisa es la solución emerge. *Precisión*, en este trabajo, es un atributo relacionado con la variación de la función objetivo al agregar un nuevo escenario, de modo que soluciones precisas son aquellas que dicha variación es pequeña en comparación con el valor de la función objetivo.

Con ese propósito un análisis de Post Optimalidad a través de la técnica de los contaminantes da la posibilidad de encontrar una cota superior e inferior para el valor de la función objetivo de un problema de programación estocástica, lo que, para este caso significa el costo total de la política de expansión [22].

Paralelamente, en otros campos de investigación, particularmente, en teoría de la información para el aprendizaje (ITL por su acrónimo en inglés) se han desarrollado técnicas que permiten obtener información a partir de distintas muestras de un fenómeno [23]. Esta teoría ha desarrollado distintos instrumentos entre lo que se destaca el hecho de poder estimar funciones de densidad de probabilidad a partir de ventanas de Parzen. Otras funciones, como la Divergencia de Cauchy Schwarz, que permite la comparación entre distribuciones. Por otra parte, Información Mutua permite medir el grado de independencia de dos variables aleatorias. Estas herramientas pueden ser usadas para hacer selección de características, comparar distribuciones de probabilidad o hacer un *cluster* de muestras tomadas a partir de realizaciones de variables aleatorias.

La gran ventaja de los enfoques usados en ITL en comparación a otros métodos como mínimos cuadrados o técnicas de *clusters* basados en nociones de distancias como *k-means*, es que no existe la necesidad de supuestos basados en algún tipo de distribución de las variables aleatorias (Gaussianidad, por ejemplo), ganando así generalidad en sus aplicaciones.

1.1. Hipótesis de la propuesta

La presente tesis busca demostrar que es posible desarrollar una metodología que permita encontrar la participación de cada tecnología (con especial atención en sistemas de almacenamiento) en el mix óptimo de generación y almacenamiento considerando las estocasticidades propias de los modelos de planificación de un sistema eléctrico de potencia. Esta metodología deberá ser eficiente en el uso de los recursos computacionales, en el sentido de ocupar una cantidad reducida de escenarios que definen la aleatoriedad del modelo.

1.2. Objetivos y alcances

La presente tesis tiene como objetivo general:

- Proponer un esquema de reducción de escenarios aplicable a problemas PGA con alta penetración de energías renovables y flexibilidad que otorgan los sistemas de almacenamiento. Así obtener un mix óptimo respecto a la incertidumbre de las FEV.

Como objetivos específicos se encuentran:

- Hacer uso de instrumentos de la Teoría de la información para el aprendizaje de máquinas en el contexto de problemas PGA. En particular hacer uso de la Información Mutua para seleccionar aquellas características que recogen de mejor manera la información de la variable aleatoria.

- Realizar una reducción de escenarios basada en técnicas de *clustering*, cuya función objetivo de cuenta de la cercanía de la distribución de probabilidad de los escenarios originales y su versión reducida.
- Proponer y evaluar expresiones que permitan determinar qué tan sensible a los escenarios es la capacidad instalada en alguna tecnología, especialmente el almacenamiento.
- Verificar los alcances de la propuesta en casos de estudio que permitan verificar las potencialidades y limitantes de cada fase. Finalmente, en un problema con dimensiones reales el cual consiste en las proyecciones del SIN Chileno para el año 2030 considerando la disponibilidad del viento como variable aleatoria.
- Comparar los resultados obtenidos con otras metodologías de la literatura.

1.3. Resultados esperados

Entre los resultados que se espera obtener es:

- Verificar que las métricas de precisión presenten un comportamiento más favorable en la medida que la cantidad de escenarios aumente. Sin embargo, se espera reconocer distintas zonas en la convergencia de estos indicadores. De modo que la mejora en los índices sea mucho mayor en aquellas zonas donde se tiene poca información.
- La mínima cantidad de escenarios que cumple con un nivel de precisión y exactitud debe depender de que tan “incierto” es la variable aleatoria. De modo que variables muy inciertas requieren una mayor cantidad de escenarios.
- La mínima cantidad de escenarios que cumple con un nivel de precisión y exactitud debe depender de que tan “sensible” es el atributo que se quiere estimar a la variable aleatoria. Atributos muy sensibles, debiesen requerir una mayor cantidad de escenarios para estimar su valor.
- La carga computacional con los escenarios reducidos debe ser mucho menor que resolver el problema con una gran cantidad de escenarios totales, la cual en muchos casos es no factible dado la capacidad de cómputo.

1.4. Estructura del documento

El documento está estructurado en cuatro capítulos, iniciando por la presente introducción. El Capítulo 2 detalla el marco teórico en el cual se basa la metodología. El Capítulo 3 por su parte, detalla la metodología propuesta, la cual consta de cuatro fases secuenciales. El Capítulo 4 presenta los dos casos de estudio y la procedencia de los datos que se usarán en el Capítulo 5, donde se aplica la metodología propuesta con el objetivo de analizar el potencial y debilidades de la metodología. Finalmente, en el Capítulo 6 se presentan las conclusiones de la tesis y se proponen futuros estudios que podrían encontrar nuevas potencialidades de la metodología.

2. Marco Teórico

El desarrollo de las tecnologías de generación a base de FEV supone que las técnicas de optimización para problemas de planificación u operación requieren que estas sean más sofisticadas por varios motivos:

Primero, las FEV tienen la característica de cambiar su nivel de generación en pequeñas ventanas de tiempo (de ahí su nombre de variable). Este hecho obliga al sistema a tener la capacidad de rampa y reservas necesarias para hacer un seguimiento de la carga que mantenga la frecuencia en niveles aceptables [24]. Dichas características se traducen en que incluir las restricciones de corto plazo y *unit commitment* en el problema de planificación significarían cambios significativos en el costo de operar el sistema [6]

Del mismo modo, las FEV poseen incertidumbre en su generación. Esto significa que al realizar una predicción de la disponibilidad de las FEV existe un error con una cierta distribución de probabilidad en dicha estimación. Si bien este error es una variable aleatoria continua, normalmente para incluir la incertidumbre endógenamente en un problema de planificación se plantean escenarios discretos con una cierta probabilidad de ocurrencia [22] De este modo, el problema de planificación tomaría en cuenta las estocasticidades que rigen el estudio.

En este punto se plantea un dilema, por una parte, una gran cantidad de escenarios asegura una representación fiel de la variabilidad y la incertidumbre de la variable aleatoria en cuestión. Sin embargo, la capacidad computacional de un problema está limitada por el número de escenarios, de modo que comúnmente esta cantidad restringe las posibilidades de resolver el problema.

En este capítulo primeramente se describe un problema de Planificación de la Generación-Almacenamiento (PGA). Para posteriormente describir distintas técnicas desarrolladas en el área de la teoría de la información para el aprendizaje de máquinas, que serán usadas para reducir escenarios conservando la información contenida en cada una de las muestras de la variable aleatoria. Finalmente, se procede a describir la técnica de post optimalidad.

2.1. Planificación de sistemas eléctricos con almacenamiento: Problemáticas y desafíos

En el contexto de planificación de sistemas eléctricos, un PGA se entiende como un problema de decidir el programa de inversiones y políticas de operación en el sector generación-almacenamiento del sistema eléctrico, con la premisa de que esto sea a un mínimo costo.

Si bien esta es una tarea clásica dentro de los sistemas de potencia, ha evolucionado con los años. En sistemas con tecnologías térmicas convencionales, donde el control sobre las tecnologías es amplio, se suele usar un enfoque con la curva de duración de carga, la cual ordena la demanda de mayor a menor, de modo que su interpretación es el número de horas que la demanda estará por sobre un determinado valor. Este tipo de problemas se puede solucionar con el uso de *screening curves* [25], encontrando la cantidad a invertir en cada tecnología y entre que rangos de demanda una tecnología sería la marginal económicamente hablando.

Sin embargo, al realizar un ordenamiento decreciente de la demanda implica que se pierda la cronología del proceso de planificación, y por tanto agrega una complicación adicional el modelar servicios complementarios que son necesarios cuando el sistema tiende a tener una alta penetración de generación con FEV. En ese contexto es que la planificación se vuelve en un problema de varias etapas que no solamente se ve acoplado por las variables de inversión, sino que además por las variables de operación. Este hecho se acentúa cuando se incorporan tecnologías que tienen como

variable de estado la energía almacenada (Embalses y sistemas de almacenamiento), pues dicha variable evoluciona etapa a etapa.

Este acople temporal remarca la complejidad del modelo, pues existe un complemento entre la resolución temporal del modelo y la calidad de los resultados de la planificación. Por tanto, variados esfuerzos se centran en desacoplar las etapas de decisión (desde la inversión hasta las etapas de operación propiamente tal). Ejemplo de aquellos son la descomposición de Benders, Dantzig-Wolfe [6], [26], o algoritmos de creación de columnas y restricciones. Incluso algoritmos heurísticos como Progressive-Hedging pueden apuntar en esa dirección.

Otro aspecto que resulta relevante de considerar es la naturaleza estocástica de la planificación. Las fluctuaciones de los precios de los combustibles, la posible baja en los costos de inversión en algunas tecnologías y, por cierto, los perfiles de la generación a partir de las FEV hacen que ciertos atributos del mix óptimo dependan fuertemente de los escenarios con los que el modelo es ejecutado. Esto hace que la planificación de un sistema forme parte de los problemas denominados de programación estocástica. Es este contexto que abre la posibilidad de formulaciones del modelo que puedan tomar en cuenta métricas de riesgos, ejemplos de lo anterior son formulaciones con restricciones denominadas de *conditional value at risk (CVaR)*, o que minimicen el valor esperado de los costos (aversa al riesgo).

Cabe señalar, que esta tesis se enfoca en planificación de la generación con sistemas de almacenamiento, sin considerar los efectos que imponen los sistemas de transmisión. Sin embargo, en muchos casos esta simplificación resulta ser fuerte, puesto que el sistema de transmisión resulta ser demandante en términos de flexibilidad. En términos del modelo, la transmisión aporta un conjunto de ecuaciones que dan cuenta por el flujo de las líneas, y su versión lineal corresponde al bien conocido flujo DC. Sin embargo, versiones más complejas pueden representar el flujo de reactivos por las líneas [4].

2.1.1. Notación de la sección

Parámetros:

$\xi_{[t]}$	Da cuenta de la variable aleatoria en el tiempo t .
$\rho_{t,\xi_{[t-1]}}[Z]$	Funcion objetivo de la etapa t , teniendo en cuenta la aleatoriedad $\xi_{[t-1]}$ y la decisión Z .
λ_t	Aversión al riesgo en el tiempo t . Es un número perteneciente al intervalo $[0,1]$
$E[Z \xi_{[t-1]}]$	Función esperanza de los costos dado que se tomó la decisión Z y la aleatoriedad esta dada por $\xi_{[t-1]}$
α_t	Tamaño de la cola de los peores escenarios, sobre los cuales se considera un inversión riesgosa.
$CVaR_{\alpha_t}[Z \xi_{[t-1]}]$	Valor condicional al riesgo, sobre los costos α_t más caros cuando la decisión de inversión es Z y la aleatoriedad está dada por $\xi_{[t-1]}$.
A_i	Anualidad de inversion en la tecnología de inversión indexada a i . [USD/kW].
A_s	Anualidad de la inversión en potencia de la tecnología de almacenamiento indexada a s . [USD/kW].
CE_s	Anualidad de la inversión en la capacidad energética de la tecnología de almacenamiento indexada a s . [USD/kWh].
$p_{\vec{k}}$	Probabilidad de ocurrencia de los escenarios dados por \vec{k} .

$CV_{i,\vec{k}}$	Costos variables de la tecnología de generación i en los escenarios \vec{k} . [USD/MWh]
CV_s	Costos variables de la tecnología de almacenamiento s . [USD/MWh]
CF	Costos de no suministro. [USD/MWh].
CU_i	Costos de encendido de la tecnología de generación i . [USD]
CD_i	Costos de apagado de la tecnología de generación i [USD].
CU_s	Costos de encendido de la tecnología de almacenamiento s . [USD]
CD_s	Costos de apagado de la tecnología de almacenamiento s . [USD].
D_{max}	Demanda máxima en el período de evaluación. [MW]
$f_{suficiencia}$	Factor de suficiencia.
$D_{t,\vec{k}}$	Carga en el tiempo t en el escenario \vec{k} . [MWh]
$\eta_{s,ch}$	Eficiencia de carga de la tecnología de almacenamiento s .
$\eta_{s,dc}$	Eficiencia de descarga de la tecnología de almacenamiento s .
$\alpha_{s,t,\vec{k}}$	Afluentes hacia la tecnología de almacenamiento en el tiempo t , en el escenario \vec{k} . [MWh]
\overline{P}_i^{max}	Máxima capacidad instalable de la tecnología de generación i . [MW]
\overline{P}_s^{max}	Máxima capacidad instalable de la tecnología de almacenamiento s . [MW]
\overline{E}_s^{max}	Máxima capacidad energética de la tecnología de almacenamiento s . [MWh].
$X_{up,i}^{min}$	Mínimo tiempo de encendido de la tecnología de generación i . [Horas]
$X_{down,i}^{min}$	Mínimo tiempo de apagado de la tecnología de generación i . [Horas]
P_i^{max}	Máxima energía generable por una unidad de la tecnología de generación i . [MW]
P_i^{min}	Mínima energía generable por una unidad de la tecnología de generación i . [MW]
M_s	Constante de linealización asociada al almacenamiento s .
fr	Proporción de la demanda que induce a un requerimiento de reserva.
$f d_{i,t,\vec{k}}$	Factor de disponibilidad de las FEV.
$Ramp_{up_i}$	Rampa de subida de la tecnología de generación i . [MWh]
$Ramp_{down_i}$	Rampa de bajada de la tecnología de generación i . [MWh]
$Ramp_{up_s}$	Rampa de subida de la tecnología de almacenamiento s . [MWh]
$Ramp_{down_s}$	Rampa de bajada de la tecnología de almacenamiento s . [MWh]
Variables:	
Z	Vector de decisiones.
P_i	Potencia instalada en la tecnología de generación i . [MW]
P_s	Potencia instalada en la tecnología de almacenamiento s . [MW].
E_s	Capacidad de energía en la tecnología de almacenamiento s . [kWh]
$G_{i,t,\vec{k}}$	Energía de la tecnología de generación i , en el tiempo t , en el escenario \vec{k} . [MWh].
$G_{s,t,\vec{k}}^+$	Energía inyectada al sistema por la tecnología de almacenamiento s , en el tiempo t , y en el escenario \vec{k} . [MWh]
$G_{s,t,\vec{k}}^-$	Energía almacenada desde el sistema por la tecnología de almacenamiento s , en el tiempo t y el escenario \vec{k} . [MWh]

$G_{F,t,\vec{k}}$	Energía no suministrada en el tiempo t en el escenario \vec{k} . [MWh]
$U_{i,t,\vec{k}}$	Número de unidades de la tecnología de generación i que se encienden en el tiempo t y en el escenario \vec{k} . Variables enteras.
$D_{i,t,\vec{k}}$	Número de unidades de la tecnología de generación i que se apagan en el tiempo t y en el escenario \vec{k} . Variables enteras.
$U_{s,t,\vec{k}}$	Número de unidades de la tecnología de almacenamiento s que se encienden en el tiempo t y en el escenario \vec{k} . Variables enteras.
$D_{s,t,\vec{k}}$	Número de unidades de la tecnología de almacenamiento s que se apagan en el tiempo t y en el escenario \vec{k} . Variables enteras.
$E_{s,t,\vec{k}}$	Estado de carga de la tecnología de almacenamiento s , en el tiempo t en el escenario \vec{k} . [MWh].
n_i	Cantidad de unidades que se instalan en la tecnología de generación i . Variables enteras.
n_s	Cantidad de unidades que se instalan en la tecnología de almacenamiento s . Variables enteras.
$X_{i,t,\vec{k}}$	Número de unidades encendidas en el instante i , tiempo t y escenario \vec{k} . Variables enteras.
$C_{s,t,\vec{k}}$	Variable disyuntiva que indica un 1 si a tecnología de almacenamiento s está descargando en el tiempo t en el escenario \vec{k} . Y 0 en caso contrario.
$R_{i,t,\vec{k}}^+$	Reserva de subida de la tecnología de generación i , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{i,t,\vec{k}}^-$	Reserva de bajada de la tecnología de generación i , en el tiempo t , en el escenario \vec{k} . [MWh].
$R_{s,t,\vec{k}}^{+1}$	Estimacion limitada por potencia de la reserva de subida de la tecnología de almacenamieto s , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{s,t,\vec{k}}^{+2}$	Estimacion limitada por energía de la reserva de subida de la tecnología de almacenamieto s , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{s,t,\vec{k}}^+$	Reserva de subida de la tecnología de almacenamiento s , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{s,t,\vec{k}}^{-1}$	Estimacion limitada por potencia de la reserva de bajada de la tecnología de almacenamieto s , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{s,t,\vec{k}}^{-2}$	Estimacion limitada por energía de la reserva de bajada de la tecnología de almacenamieto s , en el tiempo t , en el escenario \vec{k} . [MWh]
$R_{s,t,\vec{k}}^-$	Reserva de bajada de la tecnología de almacenamiento s , en el tiempo t , en el escenario \vec{k} . [MWh]

2.1.2. Función objetivo

En un problema de planificación, la función objetivo suele reflejar los costos de inversión y operación del sistema. La formulación de estos costos depende si este modelo se considerará estocástico, de modo que el funcional a minimizar podría depender de la probabilidad de ocurrencia de cada escenario.

Otro aspecto que se debe considerar en el momento de plantear la función objetivo es si es que se quiere tener o no en cuenta métricas de riesgo. Esto se puede hacer de bastantes formas, sin

embargo, una manera usual de considerar dicho elemento en la función objetivo es otorgando un mayor peso a las colas de la distribución tal como lo muestra la ecuación (2-1) [26].

$$\rho_{t,\xi_{[t-1]}}[Z] = (1 - \lambda_t)\mathbf{E}[Z|\xi_{[t-1]}] + \lambda_t CVaR_{\alpha_t}[Z|\xi_{[t-1]}] \quad (2-1)$$

Donde el primer sumando de la expresión del lado derecho da cuenta de la parte neutra al riesgo, mientras que el segundo sumando toma en cuenta las colas de la distribución. λ_t muestra el grado de aversión al riesgo del tomador de decisiones, y es un número que va entre 0 y 1. Entre más cerca de 1, quiere decir que la formulación es más aversa al riesgo.

En esta tesis se remite a formulaciones neutras al riesgo, por lo que se toma $\lambda_t = 0$ para todos los períodos. En ese caso los costos tienen dos componentes, una llamada de costos de inversión y la otra relacionada con los costos de operación esperados los cuales se reflejan en las ecuaciones (2-2) y (2-3) respectivamente.

$$\sum_i A_i P_i + \sum_s (A_s P_s + C E_s E_s) \quad (2-2)$$

$$\sum_t \sum_{\vec{k}} p_{\vec{k}} \cdot \left(\sum_i G_{i,t,\vec{k}} CV_{i,\vec{k}} + \sum_s CV_s (G_{s,t,\vec{k}}^+ + G_{s,t,\vec{k}}^-) + CF \cdot G_{F,t,\vec{k}} \right. \\ \left. + \sum_{i,t} (CU_i U_{i,t,\vec{k}} + CD_i D_{i,t,\vec{k}}) + \sum_{s,t} (CU_s U_{s,t,\vec{k}} + CD_s D_{s,t,\vec{k}}) \right) \quad (2-3)$$

La expresión de (2-2) refleja los costos anualizados de inversión en tecnologías convencionales y almacenamiento. Para sistemas de almacenamiento, los costos se desagregan en inversión en potencia y en el tamaño energético del almacenamiento. Esto permite distinguir desde una perspectiva económica sistemas de corto, mediano o largo plazo, de modo que el almacenamiento de corto plazo está caracterizado por costos de inversión en potencia bajos en comparación con almacenamientos de largo plazo. Lo contrario ocurre con la energía, donde los costos bajos comparativamente hablando están a favor de sistemas de mediano y largo plazo.

La ecuación (2-3) por su parte corresponde a una sumatoria sobre todos los instantes de tiempo, sobre todos los escenarios de la generación con cada una de las tecnologías convencionales y de FEV, junto con esto se consideran los costos operativos de operar con sistemas de almacenamiento, y finalmente la energía que no se suministra la cual se valoriza a un costo CF . Adicionalmente, debido a que el modelo considera restricciones de corto plazo, se consideran los costos de encendido y apagado de cada tecnología y almacenamiento.

2.1.3. Restricciones del problema de planificación considerando limitantes de corto plazo

Las restricciones de un problema de planificación deben considerar que las decisiones de inversión sean suficientes para abastecer la demanda de punta ponderada por un denominado factor de suficiencia. Además, se debe asegurar que la demanda sea abastecida en todas las horas (o en su defecto, considerar una potencia no suministrada, valorizada a un denominado costo de falla CF). Ambas condiciones se ven reflejadas en las ecuaciones (2-4) y (2-5)

$$\sum_i P_i + \sum_s P_s \geq D_{\text{máx}f_{\text{suficiencia}}} \quad (2-4)$$

$$\sum_i G_{i,t,\vec{k}} + \sum_s G_{s,t,\vec{k}}^+ + G_{F,t,\vec{k}} = D_{t,\vec{k}} + \sum_s G_{s,t,\vec{k}}^- \quad \forall t, \vec{k} \quad (2-5)$$

Cabe mencionar que dichas ecuaciones no consideran el uso de un sistema de transmisión, considerando esto como una hipótesis simplificadora.

Por otro lado, el almacenamiento se rige por la ecuación (2-6) de balance energético horario:

$$E_{s,t,\vec{k}} = E_{s,t-1,\vec{k}} + G_{s,t,\vec{k}}^- \eta_{s,ch} - G_{s,t,\vec{k}}^+ \eta_{s,dc} + a_{s,t,\vec{k}} \quad \forall s, t, \vec{k} \quad (2-6)$$

Donde cabe hacer la consideración que el balance energético de la etapa inicial debe considerar una carga inicial de $E_{s,1}$. Dicha ecuación considera la opción de que el sistema de almacenamiento tenga energía entrante, como por ejemplo podrían ser afluentes en una central de bombeo.

Usualmente las decisiones de inversión están acotadas, ya sea porque la fuente que provee el recurso es una limitante para invertir (muestra de ello son los caudales por los ríos), o bien, porque ciertas políticas tomadas evitan excesivas inversiones en cierta tecnología. Dichas restricciones encuentran su expresión matemática en las ecuaciones (2-7) a (2-9).

$$P_i \leq \widehat{P_i^{\text{máx}}} \quad \forall i \quad (2-7)$$

$$P_s \leq \widehat{P_s^{\text{máx}}} \quad \forall s \quad (2-8)$$

$$E_s \leq \widehat{E_s^{\text{máx}}} \quad \forall s \quad (2-9)$$

Por otro lado, en general la inversión se hace en una cantidad discreta de unidades, dada por generadores y tamaños de calderas en cantidades determinadas. Con esto, la capacidad instalada, estará dada por las siguientes ecuaciones (2-10) – (2-12).

$$P_i = n_i P_i^{\text{max}} \quad (2-10)$$

$$P_s = n_s P_s^{\text{max}} \quad (2-11)$$

$$E_s = n_s E_s^{\text{max}} \quad (2-12)$$

Finalmente, cabe mencionar las restricciones de corto plazo. Éstas están relacionadas con el *Unit Commitment* de las unidades térmicas, junto con sus tiempos mínimos de encendido y apagado relacionados con la inercia térmica de cada una de las tecnologías. Dichas ecuaciones están entre las ecuaciones (2-13) a (2-16). Cabe mencionar que las tecnologías de almacenamiento que se modelan (baterías, centrales de bombeo) no tienen un proceso térmico que implique la necesidad de que se modelen en el predespacho.

$$X_{i,t,\vec{k}} = X_{i,t-1,\vec{k}} + U_{i,t,\vec{k}} - D_{i,t,\vec{k}} \quad \forall i, t, \vec{k} \quad (2-13)$$

$$X_{i,t,\vec{k}} \leq n_i \quad (2-14)$$

Las ecuaciones (2-13) y (2-14) muestran el cambio de estado de las tecnologías convencionales térmicas.

$$X_{i,t,\vec{k}} \geq \sum_{j=t-X_{up,i}^{\min}}^t U_{i,j,\vec{k}} \quad \forall i, t, \vec{k} \quad (2-15)$$

$$n_i - X_{i,t,\vec{k}} \geq \sum_{j=t-X_{down,i}^{\min}}^t D_{i,j,\vec{k}} \quad \forall i, t, \vec{k} \quad (2-16)$$

Las ecuaciones (2-14) y (2-15) definen los tiempos de encendido y apagado admisibles. Con esto es posible definir los límites de generación de todas las unidades en cuestión.

$$G_{i,t,\vec{k}} \leq X_{i,t,\vec{k}} P_i^{\max} \quad \forall i, t, \vec{k} \quad (2-17)$$

$$G_{i,t,\vec{k}} \geq X_{i,t,\vec{k}} P_i^{\min} \quad \forall i, t, \vec{k} \quad (2-18)$$

$$G_{s,t,\vec{k}}^+ \leq P_s \quad \forall s, t, \vec{k} \quad (2-19)$$

$$G_{s,t,\vec{k}}^- \leq P_s \quad \forall s, t, \vec{k} \quad (2-20)$$

$$G_{s,t,\vec{k}}^+ \leq M_s(X_{s,t}) \quad \forall s, t, \vec{k} \quad (2-21)$$

$$G_{s,t,\vec{k}}^+ \leq M_s C_{s,t,\vec{k}} \quad \forall s, t, \vec{k} \quad (2-22)$$

$$G_{s,t,\vec{k}}^- \leq M_s(1 - C_{s,t,\vec{k}}) \quad \forall s, t, \vec{k} \quad (2-23)$$

Las ecuaciones (2-17) a la (2-23) reflejan que la generación de cada tecnología y almacenamiento debe estar entre el mínimo técnico y la potencia nominal. Las ecuaciones (2-21) a la (2-23) dan cuenta de restricciones disyuntivas para limitar la posibilidad de que el almacenamiento se cargue y se descarguen al mismo tiempo.

La incertidumbre en la generación, ya sea por salidas forzosas de algunas unidades, o bien, por la variabilidad de las FEV requiere que el sistema tenga reservas y capacidad de toma de carga para hacer frente a tales eventos. El modelamiento de las reservas está dado por las ecuaciones (2-28) a la (2-33)

$$R_{i,t,\vec{k}}^+ \leq X_{i,t,\vec{k}} P_i^{\max} - G_{i,t,\vec{k}} \quad \forall i, t, \vec{k} \quad (2-24)$$

$$R_{i,t,\vec{k}}^- \leq G_{i,t,\vec{k}} - X_{i,t,\vec{k}} P_i^{\min} \quad \forall i, t, \vec{k} \quad (2-25)$$

$$R_{s,t,\vec{k}}^{+1} = P_s - G_{s,t,\vec{k}}^+ + G_{s,t,\vec{k}}^- \quad \forall s, t, \vec{k} \quad (2-26)$$

$$R_{s,t,\vec{k}}^{+2} = \frac{(E_{s,t,\vec{k}} - G_{s,t,\vec{k}}^+ \eta_{s,dc} + G_{s,t,\vec{k}}^- \eta_{s,ch})}{\eta_{s,dc}} \quad \forall s, t, \vec{k} \quad (2-27)$$

$$R_{s,t,\vec{k}}^+ \leq R_{s,t,\vec{k}}^{+1} \quad \forall s, t, \vec{k} \quad (2-28)$$

$$R_{s,t,\vec{k}}^+ \leq R_{s,t,\vec{k}}^{+2} \quad \forall s, t, \vec{k} \quad (2-29)$$

$$R_{s,t,\vec{k}}^{-1} = G_{s,t,\vec{k}}^+ - G_{s,t,\vec{k}}^- + P_s \quad \forall s, t, \vec{k} \quad (2-30)$$

$$R_{s,t,\vec{k}}^{-2} = \frac{(E_s^{\max} - E_{s,t,\vec{k}}) + G_{s,t,\vec{k}}^+ \eta_{s,dc} - G_{s,t,\vec{k}}^- \eta_{s,ch}}{\eta_{s,ch}} \quad \forall s, t, \vec{k} \quad (2-31)$$

$$R_{s,t,\vec{k}}^- \leq R_{s,t,\vec{k}}^{-1} \quad \forall s, t, \vec{k} \quad (2-32)$$

$$R_{s,t,\vec{k}}^- \leq R_{s,t,\vec{k}}^{-2} \quad \forall s, t, \vec{k} \quad (2-33)$$

Cabe destacar que entre las ecuaciones (2-26) y la ecuación (2-33) se definen variables auxiliares para definir la reserva del almacenamiento. Esto se hace necesario porque la reserva debe considerar la capacidad tanto de potencia como de energía disponible la unidad. De modo que la reserva efectiva corresponde a la menor de las variables auxiliares, hecho que se da tanto para la reserva de subida como de bajada.

Adicionalmente, la variabilidad de la demanda, de las FEV y posibles contingencias hace necesario que el sistema tenga una capacidad de toma de carga para poder mantener la frecuencia en niveles aceptables. Las ecuaciones que dan cuenta de la capacidad de rampa de cada tecnología se detallan entre las ecuaciones (2-34) a la (2-37)

$$G_{i,t,\vec{k}} - G_{i,t-1,\vec{k}} \leq X_{i,t,\vec{k}} Ramp_{up_i} + (X_{i,t,\vec{k}} - X_{i,t-1,\vec{k}}) P_i^{min} \quad \forall i, t, \vec{k} \quad (2-34)$$

$$G_{i,t-1,\vec{k}} - G_{i,t,\vec{k}} \leq X_{i,t,\vec{k}} Ramp_{down_i} + (X_{i,t-1,\vec{k}} - X_{i,t,\vec{k}}) P_i^{max} \quad \forall i, t, \vec{k} \quad (2-35)$$

$$G_{s,t,\vec{k}}^+ - G_{s,t,\vec{k}}^- - G_{s,t-1,\vec{k}}^+ + G_{s,t-1,\vec{k}}^- \leq Ramp_{up_s} \quad \forall s, t, \vec{k} \quad (2-36)$$

$$G_{s,t-1,\vec{k}}^+ - G_{s,t-1,\vec{k}}^- - G_{s,t,\vec{k}}^+ + G_{s,t,\vec{k}}^- \leq Ramp_{down_s} \quad \forall s, t, \vec{k} \quad (2-37)$$

Una vez definidos los niveles de reserva y rampa que provee cada unidad del sistema es posible definir los requerimientos de reserva del sistema. La reserva en este modelo se dimensiona de modo de que el sistema pueda responder ante una contingencia dada por la salida forzosa de una de las unidades, o bien por el desprendimiento intempestivo de parte de la carga.

$$\sum_i R_{i,t,\vec{k}} + \sum_s R_{s,t,\vec{k}} \geq P_i^{max} \quad \forall t, \vec{k} \quad (2-38)$$

$$\sum_i R_{i,t,\vec{k}} + \sum_s R_{s,t,\vec{k}} \geq P_s^{max} \quad \forall t, \vec{k} \quad (2-39)$$

$$\sum_i R_{i,t,\vec{k}} + \sum_s R_{s,t,\vec{k}} \geq fr D_{t,\vec{k}} \quad \forall t, \vec{k} \quad (2-40)$$

Finalmente, la generación a partir de FEV se ve afectada por los perfiles de disponibilidad de éstas, de modo que, para aquellas tecnologías, su generación debe respetar la ecuación (2-41)

$$G_{i,t,\vec{k}} \leq P_i^{max} X_{i,t,\vec{k}} fd_{i,t,\vec{k}} \quad \forall i \in \{FEV\} \subseteq I, t, \vec{k} \quad (2-41)$$

Cabe mencionar que adicionalmente, en formulaciones que dan cuenta de métricas de riesgo pueden incorporarse restricciones que limitan el valor de éstas. Por otro lado, es posible considerar restricciones que incorporen de mejor manera la dinámica de las unidades [27].

En consecuencia, esta formulación corresponde a la planificación de la generación y almacenamiento de un sistema eléctrico, con restricciones de corto plazo y el modelamiento de servicios complementarios para incorporar de mejor manera las FEV.

2.2. Métricas de información a partir de muestras

Uno de los desafíos fundamentales de la sociedad actual consiste en el manejo de grandes cantidades de datos que son tomados a partir de distintos procesos en fábricas, áreas de servicios, medicina etc. El desafío radica en que el humano busca información y no datos, por tanto, eventualmente se forma un cuello de botella en la extracción de información relevante a partir de tal magnitud de datos.

El concepto información, dependiendo del contexto tiene muchas significancias, sin embargo, en este contexto tiene una definición matemática precisa, definida por Shannon en 1948 donde sienta las bases de la teoría de la información y ha tenido un alto impacto en el diseño de sistemas de comunicación eficientes y confiables. También, ha sido relevante en la teoría de probabilidad y estadística, en particular aplicada a la física estadística dando un claro entendimiento a la naturaleza de la entropía ilustrada por Jaynes [23]

Sin embargo, en todos esos estudios hasta entonces se supone un conocimiento de las distribuciones de probabilidad, situación que no se da en la mayoría de los casos y que, por tanto, en principio reduce la aplicabilidad de esta teoría. Sin embargo, investigaciones más contemporáneas proponen maneras de estimar distribuciones de probabilidad. Ventanas de Parzen es un ejemplo de tales estimaciones, lo que abre la posibilidad de medir la información contenida en un conjunto de datos, junto con otras medidas del mundo de teoría de la información [23].

En lo sucesivo, se introducen los conceptos que serán usados en la metodología para realizar una reducción de escenarios.

2.2.1. Estimación de densidades de probabilidad mediante ventanas de Parzen

Los métodos de estimación de densidades de probabilidad se pueden clasificar en métodos paramétricos y no paramétricos. Los primeros se basan en que existe un modelo subyacente para el fenómeno en cuestión, por tanto, dichos métodos buscan encontrar un valor para aquellos parámetros que determinan el modelo. Es entonces que la aplicabilidad de estos métodos se centra en el supuesto de conocer alguna expresión para la densidad de probabilidad.

Sin embargo, en contextos en donde se desconoce el modelo que regiría el fenómeno en estudio, es posible aplicar modelos no paramétricos. Unos de estos modelos son aquellos que se basan en el uso de *kernels*. Un *kernel* es una función que se define desde el dominio del fenómeno en cuestión y mapea hacia una densidad de probabilidad (es decir, mapea a \mathbb{R}), de modo que el *kernel* integrado en todo su dominio debe dar la probabilidad 1. Otro aspecto a considerarse es el ancho que tendrá el *kernel*, el cual es un parámetro

En la misma línea, dado un conjunto X_1, X_2, \dots, X_n , de muestras independientes e idénticamente distribuidas a partir de la variable aleatoria X (desconocida), un ancho del *kernel* h , entonces, la estimación de la densidad de probabilidad está dado por la expresión (2-42).

$$\hat{f}_X(z) = \sum_{i=1}^n \frac{1}{hn} K\left(\frac{z - X_i}{h}\right) \quad (2-42)$$

Para el caso particular de las ventanas de Parzen, el *kernel* tiene forma de una gaussiana, y su parámetro h , que está relacionado con el ancho del *kernel* es precisamente la desviación estándar asociada a la gaussiana.

El valor de este parámetro en general depende de heurísticas. Una de ellas es la regla de Silverman [28], la cual indica que el ancho del *kernel* está dado por la ecuación (2-43)

$$h_s = 1.06 \min \left\{ \sigma, \frac{R}{1.34} \right\} N^{-\frac{1}{5}} \quad (2-43)$$

Donde σ es la desviación estándar de las muestras X_i , y R es el intercuartil de las muestras. Cabe mencionar que la expresión anterior aplica cuando las muestras son unidimensionales. En caso de que X_i pertenezca a \mathbb{R}^n la ecuación (2-43) se ve modificada por la expresión (2-44):

$$H_S = \sigma_x \left(\frac{4}{N(2d+1)} \right)^{\frac{1}{d+4}} \quad (2-44)$$

Donde H_S es una matriz de $n \times n$. De este modo, la estimación formalmente con ventanas de Parzen con *kernel* gaussiano está dado por:

$$\hat{f}(x) = \frac{1}{N} \sum_i^N G(x - X_i, H_S^2 I) \quad (2-45)$$

$$G(x, \sigma^2 I) = \frac{1}{\sqrt{(2\pi)^d |\sigma I|}} \exp \left(-\frac{1}{2} (x - X_i) (\sigma I)^{-2} (x - X_i) \right) \quad (2-46)$$

Donde I denota la matriz identidad, d la dimensión de los datos X_i .

2.2.2. Entropía, divergencia e información mutua

Trabajar con métricas como la correlación, la desviación estándar o la varianza conlleva el supuesto inherente que las distribuciones del fenómeno en cuestión distribuyen como una Gaussiana. En cuyo caso, el fenómeno subyacente se puede caracterizar por su media y su desviación estándar, que corresponden a los dos primeros momentos de la distribución.

Sin embargo, usualmente las distribuciones no corresponden a una Gaussiana y, por tanto, se necesitan de más momentos de la distribución para caracterizar el fenómeno.

Sea, $P = (p_1, p_2, \dots, p_N)$ el conjunto de probabilidades que definen a una variable aleatoria discreta. Una medida de la entropía está dada por la ecuación (2-47).

$$H(P) = \varphi^{-1} \left(\sum_{k=1}^N p_k \varphi(I(p_k)) \right) \quad (2-47)$$

Donde $\varphi(x)$ corresponde a una función de Kolmogorov-Nagumo [23], la cual es una función continua y monótona definida en los reales. Dependiendo de la función que se tome, su expansión de Taylor indicará los momentos de la distribución que tomará en cuenta la entropía, de modo de que en general se consideran más que solo los dos primeros momentos.

Para un mayor entendimiento de esta función es conveniente analizar su uso en el contexto en el cual la teoría de la información tiene mayor uso. En telecomunicaciones se dice que, si un mensaje es perfectamente conocido *a priori*, entonces la información que contiene es cero. Dicho

de otro modo, entre menos predecible es el mensaje, se dice que cada muestra del mensaje contiene más información. La entropía mide la información promedio que tiene un mensaje, y en ese contexto, es una medida de cuan incierta es la variable aleatoria.

En el caso particular en que la expresión (2-47) toma la forma $\varphi(x) = x$, y $I(x) = -\log(x)$ entonces se tiene la conocida expresión propuesta por Shannon. Sin embargo, no es la única versión de la entropía. Una de ellas es la entropía de Renyi, la cual tiene la forma de la expresión (2-48). De hecho, al tomar el límite de $\alpha \rightarrow 1$ se obtiene la entropía de Shannon.

$$H_{R_\alpha} = \frac{1}{1-\alpha} \log \left(\sum_{k=1}^N p_k^\alpha \right) \quad \alpha > 0, \alpha \neq 1 \quad (2-48)$$

De manera análoga para las definiciones anteriores, se puede hacer la versión diferencial de estas. En particular para $\alpha = 2$ se tiene que lo que indica la ecuación (2-49), desde donde se puede definir adicionalmente el potencial de información, el que está dado por (2-50)

$$H_{R_2}(Y) = -\log \left(\int_{-\infty}^{\infty} f_Y^2(z) dz \right) \quad (2-49)$$

$$V_2(Y) = \int_{-\infty}^{\infty} f_Y^2(z) dz \quad (2-50)$$

El potencial de información requiere el conocimiento de la densidad de probabilidad de la variable en cuestión. Sin embargo, con una estimación de esta, vía ventanas de Parzen (2-45), es posible llegar a una expresión cerrada dada por (2-51), en donde se hace uso de la propiedad de convolución de Gaussianas. Es relevante destacar que dicha expresión permite estimar el potencial de información directamente de las muestras, sin suponer un modelo de la distribución de estas.

$$\begin{aligned} V(\{X\}) &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \int_{-\infty}^{\infty} G(z - X_i, H_s^2 I) G(z - X_j, H_s^2 I) dz \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G(X_i - X_j, 2\sigma^2 I) \end{aligned} \quad (2-51)$$

Cabe recalcar que el hecho de que se ocupe la entropía de Renyi con $\alpha = 2$ permite usar el teorema de la convolución entre gaussianas, por tanto, para las siguientes métricas, es una condición deseable tener términos cuadráticos de las distribuciones de probabilidad.

Análogamente, se puede definir la divergencia entre dos variables aleatorias. Dicha métrica mide la información adicional que tiene una variable aleatoria Y , dado que se conoce la variable aleatoria X . De una manera intuitiva, la divergencia mide el grado de similitud entre dos variables aleatorias, y si bien no todas las definiciones de divergencia definen una distancia por no cumplir la condición de simetría o desigualdad triangular, en muchos casos, divergencias pequeñas muestran cercanía de las variables aleatorias en el espacio de probabilidad.

La divergencia que se ocupa en esta tesis es derivada a partir de la desigualdad de Cauchy-Schwarz y se muestra en la expresión (2-52), donde la integración se hace en todo el dominio de la variable aleatoria.

$$I_{CS}(X, Y) = \log \left(\frac{(\int f_X^2(z) dz)(\int f_Y^2(z) dz)}{(\int f_X(z) f_Y(z) dz)^2} \right) \quad (2-52)$$

Es importante notar que dado que f_X y f_Y son funciones integrables y continuas por tramo, I_{CS} preserva la noción de identidad (es decir, $I_{CS} = 0$ si y solo si $f_X = f_Y$). La simetría es directa de comprobar y se puede demostrar que satisface la noción de desigualdad triangular. Por lo que I_{CS} satisface las condiciones para ser una norma.

Finalmente, la Información Mutua entre dos variables aleatorias mide el grado de independencia entre dos variables. Cabe recordar, que dos variables aleatorias son independientes, si la distribución conjunta, corresponde a la multiplicación de sus distribuciones marginales. Por tanto, la divergencia entre la multiplicación de las distribuciones marginales y la distribución conjunta se define como la información mutua entre ambas variables.

$$I_M(X, Y) = \log \left(\frac{(\int \int f_{XY}^2(z_1, z_2) dz_1 dz_2)(\int \int f_X^2(z_1) f_Y^2(z_2) dz_1 dz_2)}{(\int \int f_{XY}(z_1, z_2) f_X(z_1) f_Y(z_2) dz_1 dz_2)^2} \right) \quad (2-53)$$

Tal como la entropía de Renyi, mediante las ecuaciones (2-54) a la (2-58) se puede estimar la información mutua a partir de las muestras. El conjunto $Y_1 = \{y_{i1}\}$ da cuenta de las muestras de una variable aleatoria, análogamente para Y_2 . Del mismo modo $y = \{(y_{i1}, y_{i2})\}$, los pares de muestras i de cada variable aleatoria.

$$V(y) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left(\prod_{l=1}^2 G(y_{il} - y_{jl}, 2H_s^2 I) \right) \quad (2-54)$$

$$V^l(y_j, \{y_l\}) = \frac{1}{N} \sum_{i=1}^N G(y_{jl} - y_{il}, 2H_s^2 I), \quad l = 1, 2 \quad (2-55)$$

$$V^l(\{y_l\}) = \frac{1}{N} \sum_{j=1}^N V^l(y_j, \{y_l\}), \quad l = 1, 2 \quad (2-56)$$

$$V_{nc}(y) = \frac{1}{N} \sum_{j=1}^N \left(\prod_{l=1}^2 V^l(y_j, \{y_l\}) \right) \quad (2-57)$$

$$I_M(Y_1, Y_2, y) = \log \left(\frac{V(y) V^1(\{Y_1\}) V^2(\{Y_2\})}{V_{nc}(y)^2} \right) \quad (2-58)$$

2.3. Programación estocástica, análisis de post optimalidad

Dada la incertidumbre presente en muchos de los problemas de decisión resulta de interés tener una forma sistemática de enfrentar dichas decisiones. Lo anterior se logra matemáticamente con la denominada programación estocástica que consiste en un problema de optimización, donde la aleatoriedad de los escenarios es incorporada endógenamente. Esto hace que el problema se vuelva dependiente de los escenarios con los cuales es alimentado el modelo, y nace la pregunta de qué

tan fuerte es la dependencia. Para esto último está la teoría de post optimalidad la cual es capaz de entregar cotas entre las cuales es capaz de moverse la función objetivo.

A continuación, se presenta una descripción más acabada de ambos temas.

2.3.1. Programación estocástica

Muchos métodos de solución estocásticos aplicados a mercados financieros y energéticos se basan en aproximar una distribución de probabilidad P (que en principio corresponde a una variable aleatoria continua) por una distribución discreta con escenarios finitos [15]. En el caso de una programación multietapas, los modelos que incorporan los escenarios deben ser capaces de satisfacer el principio de no anticipación en las decisiones, el cual consiste en que las decisiones x_t en la etapa t dependen exclusivamente de las decisiones anteriores $\{x_{t-1}, \dots, x_1\}$ y de las observaciones de la variable aleatoria $\{\omega_t, \omega_{t-1}, \dots, \omega_1\}$ hasta la etapa en cuestión. Más formalmente, un problema de programación estocástica lineal se puede formular de la siguiente forma:

$$\begin{aligned} \min \quad & c_1^T x_1 + E_{\omega_1} \{\phi_1(x_1, \omega_1)\} \\ \text{s.a.} \quad & A_1 x_1 = b_1 \\ & l_1 \leq x_1 \leq u_1 \end{aligned} \quad (2-59)$$

Donde las funciones $\phi_{t-1}, t = 2, \dots, T$, están definidas recursivamente como:

$$\phi_{t-1}(x_{t-1}, \omega_{t-1}) = \inf [c_t(\omega_{t-1})^T x_t + E_{\omega_t} \{\phi_t(x_t, \omega_t)\}] \quad (2-60)$$

$$\begin{aligned} \text{s.a.} \quad & B(\omega_{t-1})x_{t-1} + A_t(\omega_{t-1})x_t = b_t(\omega_{t-1}) \\ & l_t \leq x_t \leq u_t \end{aligned} \quad (2-61)$$

Y $\phi_T = 0$. El principio de no anticipación queda determinado por las restricciones de los subproblemas ϕ_{t-1} en las ecuaciones (2-61), puesto que ligan las decisiones con los instantes anteriores y observaciones de la variable aleatoria hasta el instante t .

En este contexto, ω_{t-1} es un vector aleatorio que define los coeficientes de b_t, c_t, A_t y B_t del problema de decisión estocástico en el instante $t = 2, \dots, T$, y P corresponde a la distribución conjunta de $\omega := \{\omega_1, \dots, \omega_{T-1}\}$. Mientras que, l_t y u_t son las cotas superior e inferior para el vector de decisión x_t . Finalmente, los coeficientes de la primera etapa están fijos, pues llevan el supuesto de que se conoce el estado actual de la variable aleatoria b_1, c_1 y A_1 .

Debido a que por lo general se tiene un conocimiento parcial sobre P , y también porque numéricamente se haría más complejo trabajar con modelos para las distribuciones (por ejemplo, modelos Gaussianos), es que usualmente se trabaja en K_T escenarios para las t etapas. De modo que el problema definido por las ecuaciones (2-59) a la (2-61) queda definido por las siguientes expresiones (2-62) y (2-63).

$$\min c_1^T x_1 + \sum_{k_2=2}^{K_2} p_{k_2} c'_{k_2} x_{k_2} + \sum_{k_3=K_2+1}^{K_3} p_{k_3} c'_{k_3} x_{k_3} + \dots + \sum_{k_T=K_{T-1}+1}^{K_T} p_{k_T} c'_{k_T} x_{k_T} \quad (2-62)$$

Sujeto a las restricciones:

$$\begin{aligned}
A_1 x_1 &= b_1, \\
B_{k_2} x_1 + A_{k_2} x_{k_2} &= b_{k_2}, \quad k_2 = 2, \dots, K_2 \\
B_{k_3} x_{a(k_3)} + A_{k_3} x_{k_3} &= b_{k_3}, \quad k_3 = K_2 + 1, \dots, K_3 \\
B_{K_T} x_{a(k_T)} + A_{K_T} x_{k_T} &= b_{K_T}, \quad k_T = K_{T-1} + 1, \dots, K_T \\
l_t \leq x_{k_t} &\leq u_t, \quad k_t = K_{t-1} + 1, \dots, K_t
\end{aligned} \tag{2-63}$$

Donde $a(k_t)$ denota el instante anterior a k_t , por ejemplo, $a(k_2) = k_1$, para $k_2 = 2, \dots, K_2$. Los escenarios están dados por las secuencias de los valores aleatorios que toman los coeficientes $(c_{k_t}, A_{k_t}, B_{k_t}, b_{k_t})$. Cada subsecuencia tiene asociado un valor p_{k_t} que determina la probabilidad de el escenario k_t en su etapa correspondiente, por lo que $p_{k_t} > 0$, y además $\sum_{K_{t-1}+1}^{K_t} p_{k_t} = 1$. Lo anterior debe cumplirse para los instantes $t = 2, \dots, T$

En general, el problema anterior puede ser visto como un grafo en que cada nodo representa una decisión en un instante en particular. Las aristas corresponden a transiciones entre los distintos escenarios, de este modo un problema multietapa puede ser visto como un árbol, un abanico o una red [20]. La Ilustración 3 describe los tres esquemas de modelación, en donde se distingue un árbol porque cada instante de decisión enfrenta un futuro con incertidumbre, mientras que un abanico solo en el instante inicial presenta incertidumbre (en este caso particular se puede formular como un problema de decisión de dos etapas), mientras que las demás etapas presentan decisiones determinísticas. Finalmente, la red de escenarios es una estructura más compleja, porque existe más de un camino a un escenario en particular.

Es importante notar que todas las estructuras de problemas estocásticas pueden ser modeladas como un problema de tipo abanico. Lo anterior se logra teniendo una cantidad de escenarios que da cuenta de todos los posibles caminos a un nodo de la etapa final en el tiempo T .

Finalmente, cabe mencionar que bajo ciertos supuestos [15] la función del problema de optimización de la ecuación (2-59) corresponde a una función Lipschitz. Lo cual significa que si $\eta(\omega)$ corresponde al valor de la función objetivo en el óptimo del problema de (2-59), entonces se tiene la propiedad de la ecuación (2-64), que indica la existencia de una constante $L > 0$. En la misma fuente mencionada se indican los tipos de norma, que en general dependen del tipo de modelación elegida, donde una norma usual corresponde a la distancia de Kantorovic.

$$\left| \eta(\omega) - \eta(\tilde{\omega}) \right| \leq L \|\omega - \tilde{\omega}\| \tag{2-64}$$

Cabe destacar la relevancia de la propiedad anterior, puesto que indica que la diferencia entre los óptimos alcanzados con distintos conjuntos de escenarios será proporcional a la cercanía de los escenarios con respecto a una cierta métrica o norma. La condición anterior refleja el atributo que en esta tesis se denomina como exactitud.

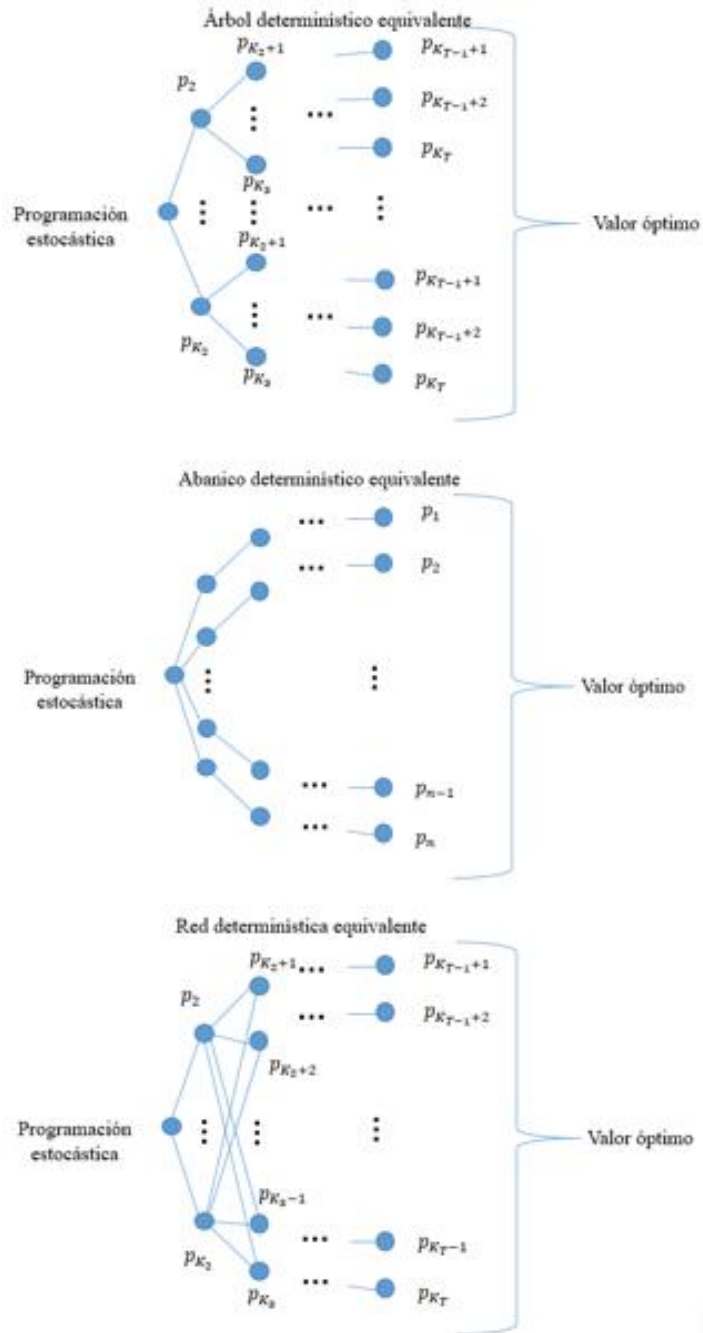


Ilustración 2-1 Representación gráfica de un problema estocástico.

2.3.2. Técnica de los contaminantes

Cualquiera sea la formulación que el problema tome, el número de escenarios resulta crucial para una adecuada representación de la variable aleatoria en cuestión y para su manejo computacional. Un gran número de escenarios asegura una buena representación de la variable aleatoria y la solución será representativa (dada la propiedad de la ecuación 2-64), sin embargo, un gran número de escenarios hace imposible de resolver el problema numéricamente debido a limitantes computacionales. De modo que es deseable trabajar con el menor número de escenarios

que converjan a una solución poco sensible a pequeños cambios en los escenarios. Sobre esto último es que se centra la teoría de post optimalidad.

De acá en más el problema de optimización se define como se expresa en la ecuación (2-65).

$$\min f(x, P), x \in \mathcal{X} \quad (2-65)$$

En donde x define las decisiones de la primera etapa, P la distribución discreta de S escenarios ω^s , asociadas a una probabilidad p_{ω^s} y \mathcal{X} a la región factible dada por las restricciones (2-63). Sea $\varphi(P)$ el valor óptimo de (2-65) y $\mathcal{X}(P)$ el conjunto de soluciones óptimas.

Se busca analizar como cambiaría el valor óptimo en el caso de agregar escenarios que distribuyen como Q (distribución contaminante). Con ese objetivo, se define la distribución P_λ como la combinación convexa entre P y Q como indica la ecuación (2-66).

$$P_\lambda = (1 - \lambda)P + \lambda Q, \quad \lambda \in (0,1) \quad (2-66)$$

Para distribuciones fijas de P y Q se define la función objetivo $f_Q(x, \lambda) := f(x, P_\lambda)$. Con esto se define $\eta_Q(\lambda)$ como el valor óptimo de $f_Q(x, \lambda)$ y $\mathcal{X}_Q(\lambda)$ como el conjunto de soluciones óptimas.

$$\begin{aligned} \eta_Q(\lambda) &= \inf_{x \in \mathcal{X}} f_Q(x, \lambda) \\ \mathcal{X}_Q(\lambda) &= \operatorname{argmin}_{x \in \mathcal{X}} f_Q(x, \lambda) \end{aligned} \quad (2-67)$$

Si se supone que $\mathcal{X}(P)$ y $\mathcal{X}_Q(1)$ son distintos de vacío, se tiene entonces que $\varphi_Q(\lambda)$ es una función cóncava en $[0,1]$, más aun, cuando $\lambda = 0$ corresponde al valor óptimo del problema (2-65), lográndose la igualdad $\varphi_Q(0) = \varphi(P)$.

Cabe recalcar la concavidad de la función $\varphi_Q(\lambda)$, por definición significa la ecuación (2-68). Dicha ecuación significa que, si se minimiza el problema estocástico con la combinación de ambas distribuciones, se obtiene un valor que será menor o igual a que si se minimiza con las distribuciones por separado.

$$\varphi_Q(\lambda) \geq (1 - \lambda)\varphi(P) + \lambda\varphi_Q(1), \quad \forall \lambda \in (0,1) \quad (2-68)$$

Por otro lado, es de interés conocer el valor marginal $\varphi'_Q(0^+)$. Bajo ciertos supuestos expuestos en [1], este valor está dado por la ecuación (2-69).

$$\varphi'_Q(0^+) = \min_{x \in \mathcal{X}(P)} [f(x, Q) - f(x, P)] = \min_{x \in \mathcal{X}(P)} f(x, Q) - \varphi(P) \quad (2-69)$$

Notar que el primer sumando del lado derecho de la expresión (2-69) en la mayoría de los casos corresponde a evaluar la función objetivo en una solución para la distribución P , puesto que el conjunto óptimo $\mathcal{X}(P)$ corresponde a un único elemento.

Con esto es posible acotar la expresión $\varphi_Q(\lambda)$ por su proyección lineal, con lo que se obtiene la cota de la expresión (2-70).

$$\varphi_Q(\lambda) \leq \varphi(P) + \lambda\varphi'_Q(0^+), \quad \forall \lambda \in (0,1) \quad (2-70)$$

La expresión (2-68) sumada a la expresión (2-70) son las denominadas cotas inferior y superior respectivamente. Es importante recalcar que no se hace ningún supuesto sobre la convexidad de las funciones $\phi_t(x_t, \omega_t)$ de la ecuación (2-59).

En el caso particular en el cual se quiere contaminar a la distribución P (inicialmente con S escenarios) con una distribución Q con S' escenarios, entonces las cotas quedan dadas por la expresión (2-71).

$$\frac{S}{S+S'}\varphi(P) + \frac{S'}{S+S'}\varphi_Q(1) \leq \varphi_Q(\lambda) \leq \frac{S}{S+S'}\varphi(P) + \frac{S'}{S+S'}f(x(P), Q) \quad (2-71)$$

Finalmente cabe mencionar que ambas cotas otorgan el atributo de soluciones precisas, debido a que el rango en el cual se moverán queda determinado por estos dos valores teóricos. Es posible ver que tomando los límites $S \gg S'$, ambas cotas convergen a $\varphi(P)$. Lo cual significa que el resultado de la programación estocástica se encuentra afecto de mayor manera a la distribución con mayor cantidad de escenarios o realizaciones.

2.4. Algoritmos de reducción de escenarios

La sección 2.3 plantea una de las problemáticas de la programación estocástica. Entre más escenarios se modelan en cada etapa, se asegura una mayor representatividad de las variables aleatorias, y por consiguiente soluciones más exactas (restricción de Lipschitz de la ecuación 2-64) y precisas por las cotas de la ecuación (2-71), sin embargo, la carga computacional se hace costosa, haciendo prohibitiva una modelación muy extensa.

Lo anterior inspira a las técnicas de reducción de escenarios, buscando modelar la estocasticidad del modelo de optimización con la menor cantidad de escenarios, pero que asegure un nivel de exactitud de la solución encontrada.

De la revisión bibliográfica se distinguen dos tipos de reducción de escenarios, aquellos que centran sus esfuerzos en tratar de representar la variable aleatoria minimizando una métrica de distancia en el espacio de probabilidades, y con esto asegurar soluciones exactas, mientras que otro conjunto de estudios busca analizar la manera en que cada escenario afecta a la solución de la programación estocástica, y en base a esto realizar la reducción de escenarios.

El primero de los casos, sienta sus bases en [29]. Inicialmente se cuenta con una distribución P con S escenarios $\omega^s := \{\omega_t^s\}_{t=1}^T$ (donde cada elemento es un vector que define las variables aleatorias de una programación estocástica de T etapas) con probabilidades $p_s > 0$ para $s = 1, \dots, S$ y $\sum_{s=1}^S p_s = 1$. Estos métodos buscan aproximar P por una distribución Q la cual posee una cantidad menor de escenarios.

Sea $J \subset \{1, \dots, S\}$ el conjunto de los índices eliminados con una cardinalidad determinada $\#J$. Luego, los escenarios restantes consisten en $\omega^{s'}$ para $s' \in \{1, \dots, S\} \setminus J$. Por lo tanto, el número de escenarios restantes esta dado por $S' = S - \#J$.

Para un conjunto J dado, la mínima distancia entre Q y P está dado por la ecuación (2-72).

$$D_K(P, Q) = \sum_{s \in J} p_s \min_{s' \notin J} c(\omega^s, \omega^{s'}) \quad (2-72)$$

Donde $c(\omega^s, \omega^{s'}) = \sum_{t=1}^T |\omega_t^s - \omega_t^{s'}|$ mide la distancia entre los escenarios ω^s y $\omega^{s'}$ y $|\cdot|$ corresponde a una norma en \mathbb{R}^n .

Las probabilidades $q_{s'}$ de los escenarios que se mantienen siguen la regla de distribución de probabilidad como sigue en (2-73), (2-74) y (2-75) las cuales se tienen que cumplir para todos los escenarios s en J

$$q_{s'} := p_{s'} + \sum_{s \in J(s)} p_s \quad (2-73)$$

$$J(s) := \{s' \in J : s' = j(s)\} \quad (2-74)$$

$$j(s) \in \arg \min_{s' \in J} c(\omega^s, \omega^{s'}) \quad (2-75)$$

Por lo tanto, cuando solo se asigna una cardinalidad $\#J$, el problema de reducción de escenarios puede ser formulado como sigue en la ecuación (2-76)

$$\min \left\{ \sum_{s \in J} p_s \min_{s' \notin J} c(\omega^s, \omega^{s'}) : J \subset \{1, \dots, S\}, \#J = S - S' \right\} \quad (2-76)$$

El cual es un problema combinatorial y es de la familia de problemas *NP-hard*, lo cual da cuenta del grado de dificultad en resolver dicho problema, por ende, se suelen usar heurísticas. Las más usadas son las llamadas *backward-reduction* y *forward-selection* [16].

Backward-reduction se basa en la regla de eliminar un único elemento, por lo que el problema (2-76) queda como expresa la siguiente ecuación.

$$\min_{k \in \{1, \dots, S\}} p_k \min_{s' \neq k} c(\omega^k, \omega^{s'}) \quad (2-77)$$

Luego el algoritmo continúa recursivamente eliminando escenarios sobre el conjunto de escenarios resultantes hasta quedar con los escenarios deseados inicialmente.

De manera análoga, el algoritmo *forward-selection*, busca que solo un escenario se mantenga, en ese caso el problema (2-76) queda como sigue:

$$\min_{u \in \{1, \dots, S\}} \sum_{s=1}^S p_s C(\omega^s, \omega^u) \quad (2-78)$$

Al igual que en *backward-reduction*, el algoritmo continúa aumentando el conjunto de escenarios preservados hasta que se igualan la cardinalidad de los elementos.

Ambos algoritmos no dan garantías de optimalidad de la reducción de escenarios (tal como lo propone 2-76) y como alternativas se han propuesto otros algoritmos que muestran mejores desempeños con respecto a los tiempos de cómputos como *simultaneous backward reduction*, y *fast forward selection* [16]. Incluso, la propuesta central de [20] consiste en una reducción de escenarios mediante una optimización por enjambre de partículas.

Todas estas propuestas se basan en tratar de minimizar un criterio de distancia (típicamente la distancia de Kantorovic), los cuales son criterios probabilísticos, y no consideran el problema de

optimización al cual se someten. Esto abre la puerta a la segunda familia de algoritmos de reducción de escenarios, los cuales buscan tomar en cuenta la manera en que cada escenario afecta a la programación estocástica.

En [21] se propone el algoritmo llamado *forward selection in recourse clusters (FSRC)* que trabaja con índices de sensibilidad que dan cuenta de cómo afecta el escenario a la solución del problema estocástico. Un índice de sensibilidad puede ser cualquier atributo medible y se sugiere no tomar atributos redundantes, puesto que aumentaría la dimensión del problema. Por ejemplo, en un problema de operación de un sistema eléctrico, existe una directa relación entre el número de generadores GNL encendidos, y la generación con dicho combustible.

Para obtener el valor de dichos índices de sensibilidad, se ejecuta el problema determinístico para los s escenarios y se obtiene el valor de dichas variables. Finalmente se procede a aplicar una técnica de clustering (*k-means*) sobre el valor de los índices de sensibilidad para posteriormente seleccionar escenarios de cada *cluster* asignando una probabilidad de manera similar a como lo indican las ecuaciones (2-73) a la (2-75)

Esta técnica no considera ningún criterio de cercanías en el espacio de probabilidad, pero si como esto afecta al problema de programación estocástico, y por tanto muestran distanciamiento respecto a la primera técnica de algoritmos.

3. Metodología

La propuesta para reducción de escenarios consiste en cuatro fases principales, las cuales se muestran en la Ilustración 4 y se describirán en detalle a lo largo del presente capítulo.

A modo de resumen, los datos de entrada indican los siguientes aspectos:

- Toda la información relevante para un estudio de PGA (costos de todas las tecnologías disponibles, parámetros técnicos. Más generalmente, todos los aspectos que son descritos en la sección 2.1).
- n escenarios que representan adecuadamente la incertidumbre del sistema (por ejemplo, incertidumbre en costos, series de viento o sol que tomen en cuenta la correlación temporal y/o espacial).
- Un conjunto de características tomadas a partir de los resultados del PGA (por ejemplo, los costos totales o la capacidad instalada en alguna tecnología en particular).

La primera fase, llamada Selección de Características, recibe n escenarios y el conjunto de características evaluadas. Esta fase tiene como objetivo determinar cuál de todas las características captura de mejor manera la información contenida en las n realizaciones de la variable aleatoria que rige el sistema. Cabe mencionar que el cálculo del valor de las características impone la necesidad de n ejecuciones de un PGA determinístico.

Por su parte, la segunda fase llamada Reducción de Escenarios considera una técnica de *clustering* de n elementos a m familias. Este *clustering* busca optimizar una métrica basada en la teoría de la información (Divergencia de Cauchy-Schwarz). Con esto, la función de densidad de probabilidad de los m *clusters* tendrá la mejor aproximación a la densidad de probabilidad original de los n escenarios. De este modo se busca que los m escenarios seleccionados sean consistentes con la noción de exactitud.

Posteriormente, la fase tres propone la posibilidad de realizar un Análisis de Post Optimalidad, centrándose en aspectos específicos del problema de planificación los cuales son definidos por el usuario, como por ejemplo, la evolución del mix de almacenamiento en un sistema de potencia. Con este propósito, la técnica de los contaminantes se aplica con el fin de encontrar una cota superior e inferior para los costos totales, dando respuestas a la noción de precisión de la solución encontrada. Además, se procede a calcular índices de adaptabilidad, que son propuestos en la sección 3.3.

Si el criterio de detención no se cumple, en la fase cuatro se propone una técnica de ajuste. En caso contrario, los resultados del PGA se consideran como resultados finales y los aspectos específicos del PGA son considerados como óptimos.

Se espera que los esfuerzos computacionales sean mucho menores, comparados con el PGA estocástico con n escenarios.

3.1. Nomenclatura del capítulo

n	Cantidad inicial de escenarios.
m	Cantidad de escenarios reducidos.
$f(x)$	Distribución de probabilidad de la característica seleccionada en la Fase I, estimada con ventanas de Parzen con n muestras iniciales.
$g(x)$	Distribución de probabilidad de los m <i>clusters</i> .
$\{a_i\}$	Valore que difinen el inicio y el término de cada <i>cluster</i> de $g(x)$.

η	Factor que multiplica el gradiente. Indica la cantidad de veces que los puntos a_i se moverán por la dirección de máximo decrecimiento.
ω_i	Ruido blanco aditivo Gaussiano en la iteración i .
UB_j	Cota superior de los costos totales al añadir el escenario ω_j .
LB_j	Cota inferior de los costos totales al añadir el escenario ω_j .
x_p	Política óptima de planificación cuando se da el escenario que distribuye como P .
$F(x_p, P)$	Costo de planificación cuando se decide seguir la política óptima x_p y con el escenario P .
R_i	Indicador de precisión i .
$Df(P, \{\omega_j\})$	Derivada direccional de los costos de planificación con respecto a la dirección $P - \{\omega_j\}$.
ΔS_i	Índices de adaptabilidad i
λ_i	Variable dual asociada a la tecnología i cuando se impone su capacidad instalada en el problema de planificación.[USD/MW]
X_i	Capacidad instalada en la tecnología i . [MW]
I_i	Información mutua entre la variable aleatoria y la capacidad instalada i obtenida en la fase I.

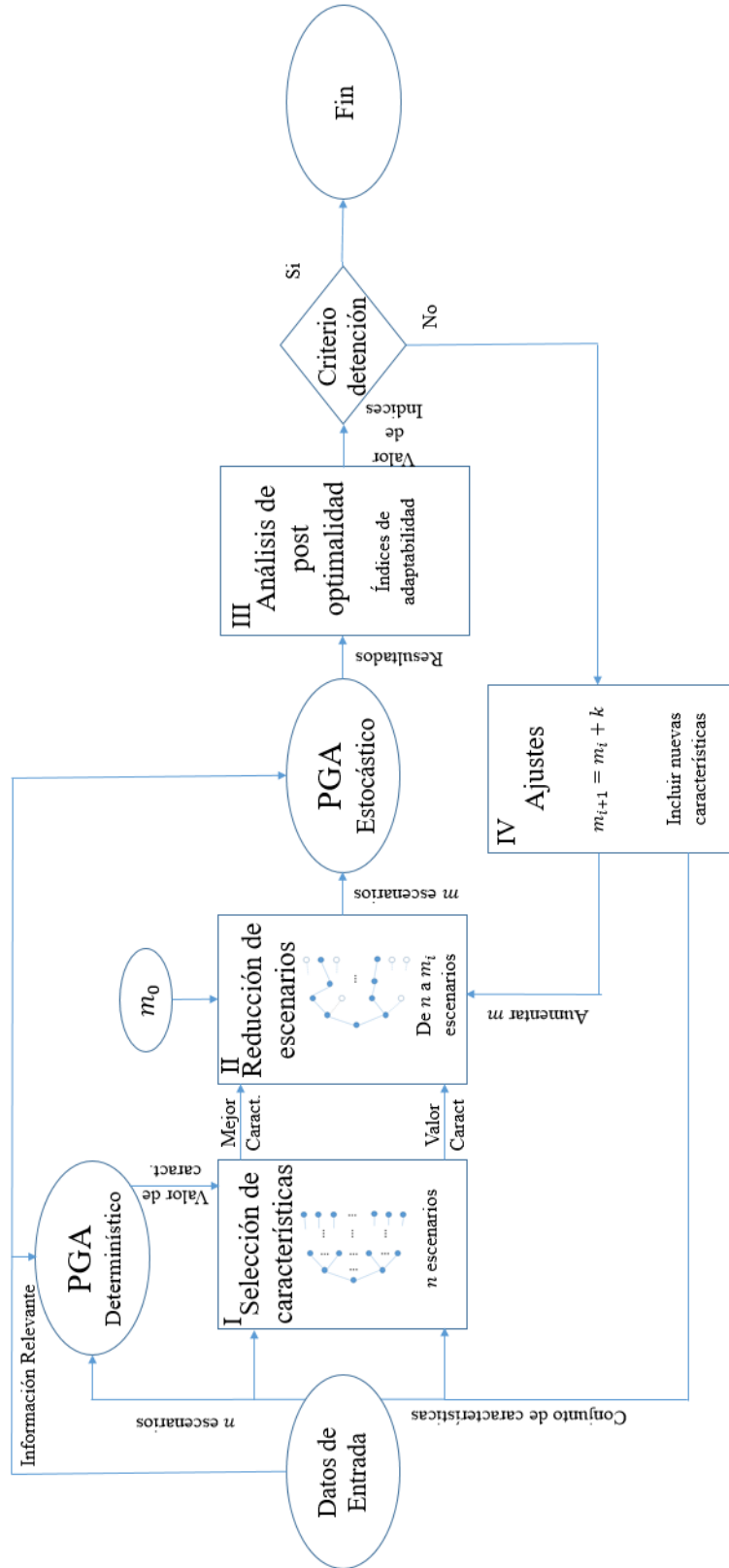


Ilustración 3-1 Diagrama de la propuesta metodológica.

3.2. Fase I: Selección de características

Cada escenario puede ser caracterizado por la forma en que éste afecta a la solución del PGA. Para medir esto, un PGA determinístico es ejecutado sobre cada uno de los n escenarios de entrada y las características son tomadas a partir de los resultados de éste. Una característica puede ser cualquier atributo medible sobre dichos resultados (por ejemplo, la capacidad instalada en almacenamiento, número de ciclos de los sistemas de almacenamiento, o los costos totales del sistema).

Cabe mencionar que el número n de escenarios sobre el cual se realizan la reducción de escenarios, es un subconjunto de una cantidad s de escenarios que describen completamente la variable aleatoria. Muchas veces la variable aleatoria como tal corresponde a una variable continua, por lo que s es una cantidad infinita no numerable. Por tanto, un supuesto en el que se basa toda la metodología es que los n escenarios corresponden a una buena aproximación de los s escenarios.

El objetivo de esta fase es escoger cuál de todas las características propuestas en los datos de entrada representa la variable aleatoria de mejor forma desde un punto de vista asociado a las métricas de probabilidad. En ese sentido, esta fase puede ser considerada como una reducción de dimensionalidad de la variable aleatoria (la cual corresponde a un vector \mathbb{R}^t) a un valor escalar, o bien, a un vector con k dimensiones ($k < t$). Por tanto, el beneficio de esta fase se ve en una disminución de cálculo en las fases sucesivas, en donde trabajar con un gran número de dimensiones hace al algoritmo vulnerable a problemas como la denominada “maldición de la dimensionalidad”.

Para este propósito se computa la información mutua (I_M) entre la característica y la variable aleatoria mediante la ecuación (2-53). De este modo aquella que presente mayor información mutua con la variable aleatoria se escoge como la característica que representará a la variable aleatoria.

Incluso, es posible explorar combinaciones de características, de modo que se calcula la I_M entre la variable aleatoria y la distribución conjunta de las características. Este enfoque debería aspirar a mayores valores de I_M a costa de tener una característica seleccionada con mayor dimensión.

Cabe recalcar que la I_M se calcula a partir de las estimaciones de la densidad de probabilidad mediante ventanas de Parzen, por lo que no existe ningún supuesto sobre el modelo de distribución de probabilidad de la variable aleatoria.

3.3. Fase II: Reducción de escenarios

Una vez electa la característica con mejor I_M , el paso siguiente corresponde a aplicar una reducción de escenarios basado en una técnica de *clustering*. Aun cuando no existe una única definición de *clustering*, en esta tesis este concepto consiste formalmente en dividir el dominio (el intervalo entre el máximo y el mínimo valor de dicha característica) de acuerdo a cierta figura de mérito.

La distribución de cada *cluster* considera que todos los elementos dentro de éste tienen la misma densidad de probabilidad. Con este propósito, el dominio es dividido en tantos subintervalos como número de m *clusters* son requeridos. Con esto, la distribución del *cluster* puede ser expresada como se muestra en la siguiente ecuación.

$$g(x) = \begin{cases} \int_{a_i}^{a_{i+1}} f(y)dy, & x \in [a_i, a_{i+1}], \quad \forall i = 1, \dots, m \\ 0, & x < a_1, x > a_{m+1} \end{cases} \quad (3-1)$$

Donde $f(x)$ corresponde a la distribución de probabilidad de la característica elegida en la Fase I. Los a_i corresponden a los puntos donde se divide el codominio y que señalan el inicio y el fin del *cluster* en cuestión.

En este caso, la figura de mérito usada corresponde a la divergencia de Cauchy-Schwarz I_{CS} (ecuación 2-52) la cual puede ser estimada a partir de la distribución de probabilidad de la variable aleatoria $f(x)$ y su versión clusterizada $g(x)$. La cual presenta la siguiente expresión:

$$I_{cs}(a_i) = \log \left(\frac{(\int f(x)^2 dx)(\int g(x)^2 dx)}{(\int g(x)f(x)dx)^2} \right) \quad (3-2)$$

Esta figura de mérito tiene como objetivo que la distribución resultante de *clusters* sea cercana en el espacio de probabilidad a la distribución original de la característica seleccionada. Esta condición es deseable porque una PGA estocástica es una función de Lipschitz respecto a la variable aleatoria. Esto, en virtud de la ecuación (2-64) implica que las diferencias entre las funciones objetivo (es decir, los costos totales de planificación según ambas distribuciones de probabilidad) son proporcionales a la distancia en el espacio de probabilidad. De modo que esta fase tiene como objetivo satisfacer la noción de exactitud de la solución.

De acuerdo con lo anterior, los a_i de la ecuación (3-1) son electos con el objetivo de minimizar I_{CS} de la ecuación (3-2). Sin embargo I_{CS} corresponde a una función no lineal, por lo que la minimización es realizada mediante la ampliamente conocida técnica del gradiente descendente, donde el gradiente de I_{CS} es dado por la siguiente ecuación:

$$\frac{\partial I_{CS}(a_i)}{\partial a_i} = \frac{1}{\int g(x)^2 dx} \left(\frac{\partial(\int g(x)^2 dx)}{\partial a_i} - \frac{\partial(\int f(x)g(x)dx)^2}{\partial a_i} \frac{(\int g(x)^2 dx)}{(\int f(x)g(x)dx)^2} \right) \quad (3-3)$$

Las derivadas parciales están dadas por las ecuaciones (3-3) y (3-4):

$$\begin{aligned} \frac{\partial(\int g(x)^2 dx)}{\partial a_i} = & \frac{2 \left(\int_{a_{i-1}}^{a_i} f(y)dy \right) f(a_i)}{a_i - a_{i-1}} - \frac{\left(\int_{a_{i-1}}^{a_i} f(y)dy \right)^2}{(a_i - a_{i-1})^2} - \frac{2 \left(\int_{a_i}^{a_{i+1}} f(y)dy \right) f(a_i)}{a_{i+1} - a_i} \\ & + \frac{\left(\int_{a_i}^{a_{i+1}} f(y)dy \right)^2}{(a_{i+1} - a_i)^2}, \forall i \in \{2, \dots, m\} \end{aligned} \quad (3-4)$$

$$\begin{aligned}
& \frac{\partial (\int f(x)g(x)dx)^2}{\partial a_i} \\
&= 2 \left(\sum_{j=1}^m \frac{(\int_{a_j}^{a_{j+1}} f(y)dy)^2}{a_{j+1} - a_j} \right) \left(\frac{2 (\int_{a_{i-1}}^{a_i} f(y)dy) f(a_i)}{a_i - a_{i-1}} \right. \\
&\quad - \frac{(\int_{a_{i-1}}^{a_i} f(y)dy)^2}{(a_i - a_{i-1})^2} + \frac{(\int_{a_i}^{a_{i+1}} f(y)dy)^2}{(a_{i+1} - a_i)^2} \\
&\quad \left. - \frac{2 (\int_{a_i}^{a_{i+1}} f(y)dy) f(a_i)}{a_{i+1} - a_i} \right) \forall i \in \{2, \dots, m\}
\end{aligned} \tag{3-5}$$

Con esto, el método del gradiente descendente se inicia con valores predeterminados a_i^1 y posteriormente son actualizados mediante la regla de actualización de la ecuación (3-6), donde el parámetro η usualmente es elegido mediante alguna heurística, ya que formalmente dicho valor sale de resolver el problema de minimización de $I_{CS}(a_i^t)$ con respecto a η , sin embargo, esto aumenta la carga computacional.

La racionalidad detrás del η indica cuánto hay que avanzar por la dirección del gradiente con el fin de minimizar el funcional en cuestión, por tanto, en regiones lejanas del óptimo se hacen convenientes η con gran valor para alcanzar el mínimo, y en caso contrario, η pequeños para tener una convergencia más fina hacia el valor mínimo.

Cualquiera sea el caso, dependiendo del η escogido, en un problema convexo el algoritmo del gradiente descendente debería converger al mínimo global (puesto que solo existe uno), mientras que, si no se tiene garantías sobre la convexidad del problema, es posible que el algoritmo converja a un mínimo local. Por dicho motivo es que la ecuación (3-6) tiene un sumando ω_i , el cual da cuenta de un ruido blanco aditivo gaussiano, el cual tiene como objetivo sacar de la zona de convergencia a un mínimo local, haciendo más factible la convergencia a un mínimo global. Por tanto, el tamaño de ese ruido (SNR por su acrónimo en inglés *Signal to Noise Ratio*) debe ser suficientemente grande para mover de mínimos locales la convergencia del algoritmo, y suficientemente pequeño para no anular la adición del gradiente.

$$a_i^t = a_i^{t-1} - \eta \nabla I_{CS}(a_i^{t-1}) + \omega_i \tag{3-6}$$

Finalmente, el escenario representante de cada *cluster* corresponde al escenario más cercano a la media del *cluster* y se le asigna la probabilidad correspondiente al intervalo del *cluster*.

3.4. Fase III: Análisis de post optimalidad

Recapitulando, una vez terminada la Fase II, se tienen m *clusters* cuya distribución aproxima a la distribución de los n escenarios originales. Como ya se mencionó anteriormente, esto con la propiedad de Lipschitz (2-64) asegura que las soluciones sean exactas. Sin embargo, no asegura nada acerca de la precisión de la solución. En ese contexto cabe preguntarse en cuanto podría cambiar la solución cuando se agrega un escenario adicional a los m ya escogidos. Si la solución es suficientemente robusta, entonces la solución debería sufrir cambios menores.

Para este propósito se hace uso del análisis de post optimalidad descrito en la sección 2.3. Es posible encontrar cotas superiores UB_j e inferiores LB_j del valor de los costos de planificación cuando se agrega un escenario ω_j que puede ser cualquiera de los $(n - m)$ escenarios restantes. Las expresiones para dichas cotas se justifican por la expresión (2-71), y su expresión concreta está determinada en las ecuaciones (3-7) y (3-8)

$$UB_j = \frac{m}{m+1}F(x_p, P) + \frac{1}{m+1}F(x_p, \{\omega_j\}) \quad (3-7)$$

$$LB_j = \frac{m}{m+1}F(x_p, P) + \frac{1}{m+1}F(x_{\{\omega_j\}}, \{\omega_j\}) \quad (3-8)$$

Donde $F(x_p, P)$ representa el costo de planificación de la PGA. Mientras que, x_p corresponde al plan de inversión óptimo del PGA con m escenarios que dan lugar a la distribución P .

La diferencia entre la máxima de las cotas superiores y la mínima de las cotas inferiores es un indicador de la precisión de la solución alcanzada por los m escenarios. El valor R_1 (indicado en la ecuación 3-9) se propone como una métrica de precisión.

$$R_1 = \frac{\left(\max_j UB_j - \min_j LB_j \right)}{\min_j LB_j} \quad (3-9)$$

Por otro lado, la derivada direccional dada por la expresión (2-70) da cuenta de la variación al agregar un escenario. Dada la linealidad en la formulación del PGA, la solución de x_p suele ser única. Por lo que la expresión (2-70) es reemplazada por la (3-10) cuando se agrega un escenario ω_j . De este modo, R_2 también da cuenta de la precisión, puesto que aquellas soluciones estables implican un cambio relativo pequeño al agregar dicho escenario.

$$Df(P, \{\omega_j\}) = f(x_p, \{\omega_j\}) - f(x_p, P) \quad (3-10)$$

$$R_2 = \frac{\mathbb{E}_j \left(Df(P, \{\omega_j\}) \right)}{f(x_p, P)} \quad (3-11)$$

Donde $\mathbb{E}_j(\cdot)$ corresponde al operador esperanza sobre todos los j pertenecientes a alguno de los $(n - m)$ escenarios restantes.

Sin embargo, en muchas ocasiones, más importante que los costos de planificación, resulta de interés saber algún aspecto particular del PGA, por ejemplo, la capacidad instalada en sistemas de almacenamiento. Con este objetivo es que se proponen los índices de adaptación.

Éstos índices se basan en el cálculo de las variables duales, los cuales se obtienen al imponer el valor en un PGA determinístico el valor del atributo en cuestión. Al calcular las variables duales, se aplica el supuesto de que el PGA debe ser convexo, por tanto, las variables enteras que hace uso un PGA con restricciones de corto plazo (sección 2.1) son relajadas linealmente. Finalmente cabe recordar que las variables duales tienen la interpretación de cuánto aumentaría el costo de planificación en caso de agregar una unidad adicional del atributo.

Para el caso particular de sistemas de almacenamiento, se proponen tres diferentes índices de adaptación mostrados en las ecuaciones (3-12) -(3-14)

$$\Delta S_1 = \frac{\left(\max_j UB_j - \min_j LB_j \right)}{\lambda_{storage}} \quad (3-12)$$

$$\Delta S_2 = \frac{\max_j UB_j - \min_j LB_j}{\sum_t \frac{\lambda_t X_t}{\sum_i X_i}} \frac{X_s}{\sum X_i} \quad (3-13)$$

$$\Delta S_3 = \frac{\max_j UB_j - \min_j LB_j}{\sum_t \frac{\lambda_t I_t}{\sum_i I_i}} \frac{I_s}{\sum I_i} \quad (3-14)$$

El primer índice ΔS_1 (3-12) considera que la máxima diferencia de costos posibles al agregar un escenario se debe única y exclusivamente a inversiones en sistemas de almacenamiento. Por lo tanto, dicho índice representa el peor de los escenarios de variación en la inversión, puesto que en general la diferencia de costos que arrojan las cotas de la ecuación (3-7) y (3-8) no se deben netamente a inversiones en este tipo de tecnologías.

Por su parte, ΔS_2 (3-13) toma como supuesto que la proporción entre las capacidades de las tecnologías se mantiene constante. Por ese motivo la diferencia máxima de costos de planificación es dividida por la inversión promedio de las tecnologías. Con esto se obtiene la potencia adicional que explicaría la diferencia de costos de las cotas, la cual se prorroga según la participación del almacenamiento en el mercado.

Finalmente, ΔS_3 (3-14) corresponde al índice de adaptación estadístico, el cual considera cuán afecto es el sistema de almacenamiento a la variable aleatoria, lo cual se mide con uso de la información mutua I_M . Es por este motivo que es necesario que en el conjunto de características que alimentan la Fase I se encuentren la capacidad instalada en cada una de las tecnologías, de modo que en dicha fase se haga el cálculo de la información mutua I_M entre cada tecnología y la variable aleatoria que rige el PGA. De esta forma, la diferencia entre cotas es dividida por el costo ponderado por la medida de cuán afecto es la inversión en dicha tecnología a cambios en la variable aleatoria. Con esto se obtiene el aumento en capacidad promedio de las tecnologías afectas por la variable aleatoria, por lo que el aumento de inversión en la tecnología de almacenamiento corresponde a un prorrateo según cuán sensible es la inversión de esta tecnología a la variable aleatoria.

3.5. Fase IV: Ajustes

Cualquiera de los índices de adaptación o de precisión pueden ser considerado como la figura de mérito del criterio de detención, es decir, cuando alguno de los índices sea menor que un cierto umbral entonces se considera que el algoritmo ha convergido a la solución que responde a los criterios de exactitud y precisión necesarios.

En caso contrario, se distinguen dos vías para mejorar dichos índices:

- Agregar nuevas características a evaluar en la Fase I que capture información relevante de la variable aleatoria que rige el PGA. Alternativamente se puede considerar una combinación de las características existentes de modo de calcular la información mutua entre la densidad de probabilidad conjunta y la variable aleatoria. Ambos criterios buscan tener una mejor representatividad de la variable aleatoria, y, por tanto, la

reducción de escenarios de la Fase II aspira converger a una probabilidad de los escenarios reducidos más cercanos en el espacio de probabilidad, mejorando la exactitud de la solución.

- Aumentar el número m a m' de escenarios reducidos ($m < m'$). Esta alternativa, también mejorará la exactitud, puesto que la densidad $g(x)$ (3-1) tendrá más grados de libertad, por lo que la divergencia de Cauchy-Schwarz llegará a un valor óptimo menor o igual al caso con m escenarios, de modo que tiene implicancias directas en la exactitud de la solución. Además, hace que las cotas superior e inferior de las ecuaciones (3-7) y (3-8) converjan a $F(x_p, P)$, de modo que todos los indicadores de precisión y adaptación propuestos tengan un menor valor, haciendo factible que se cumpla el criterio de detención. Sin embargo, a costa de esto, el PGA se hace más costoso computacionalmente debido a la incorporación de más escenarios a éste.

3.6. Comparación con algoritmos en la literatura

Tal como se menciona en la sección 2.4, se distinguen dos tipos de algoritmos de reducción de escenarios. Aquellos que se basan en minimizar distancias en espacio de probabilidad y otros que buscan realizar *clustering* usando *k-means* sobre índices de sensibilidad.

La Tabla 1 compara bajo seis aspectos las metodologías ya mencionadas y la propuesta de esta tesis, y dicha comparación se detalla en lo sucesivo.

Del mismo modo que se menciona en la sección 2.4, que el problema de reducción de escenarios óptimo desde un punto de vista de métricas en espacios de probabilidad corresponde a un problema *NP Hard*. Esto ha implicado que los algoritmos desarrollados (*forward selection*, *backward reduction* y sus derivados) corresponden a heurísticas que no dan garantías de optimalidad en la reducción de escenarios. Así mismo, los algoritmos basados en *clustering* con *k-means* son algoritmos que no apuntan directamente a la optimalidad en la reducción de escenarios, y por tanto son una heurística en sí misma. Análogamente, la propuesta metodológica de la tesis realiza un *clustering* cuya función objetivo busca minimizar una noción de distancia en el espacio de probabilidades. El proceso de optimización se opera con el gradiente de paso descendente debido a la no linealidad de I_{CS} de modo que también cae en heurísticas.

Debido a la condición de Lipschitz de la PGA, mientras se busque que las probabilidades de los escenarios originales y reducidos sean cercanas en el espacio de probabilidad, los algoritmos responderán a la noción de exactitud. Claramente los algoritmos basados en métricas de probabilidad responden inherentemente a la condición de exactitud. Los algoritmos basados en *clusters k-means* sobre índices de sensibilidad (similares a las Características en la nomenclatura de la tesis), responden a un criterio geométrico sobre índices que no necesariamente apuntan a la cercanía de las probabilidades de las variables aleatorias. Sin embargo, en la bibliografía mencionada sobre esos métodos [21], se hacen análisis de sensibilidad sobre muchos posibles índices, donde se tiene mejor desempeño en aquellos índices que representan de mejor manera la variable aleatoria que rige el PGA. La metodología propuesta en cambio, busca la característica que mejor representa a la variable aleatoria y para esto minimiza la información mutua entre la característica y la variable. Sobre ésta se realiza un *clustering* que minimiza la divergencia de Cauchy-Schwarz, de modo que busca responder a la noción de exactitud desde los objetivos de sus métricas.

	Basados en métricas de probabilidad	Basados en cluster sobre índices de sensibilidad.	Propuesta tesis
<i>Necesidad de heurísticas.</i>	Si	Si	Si
<i>Responde a noción de soluciones exactas</i>	Si	No, pero en general se consigue.	Si
<i>Responde a noción de soluciones precisas.</i>	Algunas.	No	Si
<i>Forma de evaluar el desempeño del algoritmo</i>	Respecto a los costos totales	Respecto a los costos totales	Respecto a los costos totales o atributos de interés
<i>Cantidad m de escenarios reducidos.</i>	Parámetro definido por usuario.	Parámetro definido por usuario.	Resultado del algoritmo.
<i>Tiempos computacionales en la reducción esperados</i>	Menores	Medios	Mayor a los algoritmos de la literatura.

Tabla 3-1 Comparación de la propuesta metodológica con los algoritmos de la literatura

Respecto a la noción de precisión, algunos artículos perteneciente a los que se basan en métricas de probabilidad [30] hacen uso de los análisis de post optimalidad, sin embargo, en general no se hace. Las técnicas basadas en *clustering* tampoco hacen uso de dichas técnicas. En contraste, los análisis de post optimalidad son parte esencial de la propuesta metodológica y dan origen a los denominados índices de adaptabilidad. De modo que la propuesta metodológica, además puede evaluar el desempeño de la misma respecto a atributos ajenos a la función objetivo, como por ejemplo la capacidad instalada en almacenamiento.

La cantidad m en general se propone que sea una cantidad tal que haga que la distancia entre las probabilidades sea suficientemente cercana, sin embargo, este valor debe ser cambiado por el usuario en la mayoría de los artículos publicados. La propuesta metodológica incluye un lazo cerrado para determinar la cantidad de escenarios que consiga satisfacer el criterio de detención propuesto.

Finalmente, todas las ventajas que consigue la propuesta metodológica se hacen a cambio de un mayor tiempo de cómputo debido a la necesaria ejecución de PGA determinísticos. Sin embargo, se espera que todas las propuestas tengan cargas computacionales menores al caso de resolver el PGA estocástico con n escenarios.

4. Casos de estudio

Para verificar el potencial de la metodología, se proponen dos casos de estudio, los cuales son un PGA de prueba y un PGA aplicado al sistema eléctrico nacional chileno (SEN).

El PGA de prueba corresponde a un modelo con un horizonte de 168 horas (una semana) sin restricciones de corto plazo. La variable aleatoria corresponde a 100 muestras de viento las cuales son generadas con un modelo auto regresivo de media móvil [31](ARMA por su acrónimo en inglés). Cada muestra corresponde a una serie sintética de longitud igual al horizonte de análisis del problema. El PGA aplicado al SEN por su parte, aplica un modelo de planificación con una operación equivalente a seis meses, con la realidad del SEN proyectada al año 2035. Del mismo modo, la variable aleatoria corresponde a 100 muestras de viento.

Es entonces que se tiene dos PGA que corresponden a modelos de programación estocástica de 100 escenarios, y sobre estos se aplica la metodología de reducción de escenarios.

A continuación, se describen en detalle ambos casos de estudio y los propósitos de cada uno.

Cabe recalcar que los modelos de planificación usados en estos experimentos no hacen uso del sistema de transmisión. Sin embargo, en términos de la metodología este aspecto no es una limitante, puesto que usualmente el conjunto de restricciones que impone la transmisión es un conjunto convexo, y no afecta a los supuestos necesarios para la aplicación de dicha metodología.

4.1. PGA de prueba

El objetivo de este modelo es tener un análisis acabado del funcionamiento de cada fase de la metodología, sin embargo, no se pretende tener conclusiones sobre la realidad de un mercado eléctrico, de modo que los precios y demanda no necesariamente son representativos del acontecer de algún mercado eléctrico.

El siguiente cuadro resume las tecnologías, sus costos y parámetros técnicos. Cabe mencionar que las inversiones que se incorporan al PGA son las correspondientes a la semana de evaluación.

<i>Tecnología</i>	Costos de inversión [USD/kW]	Costos de operación y mantenimiento [USD/MWh]	Máxima capacidad instalable [MW]
<i>Carbón.</i>	250	43	20.000
<i>Gas ciclo combinado.</i>	77.5	171	20.000
<i>Diesel.</i>	66.6	230	20.000
<i>Biomasa</i>	56.5	18	5.639
<i>Solar fotovoltaica.</i>	108.3	0	1.237.903
<i>Turbinas eólicas</i>	50	0	37.477
<i>Hidráulicas de pasada.</i>	75	0	20.000

Tabla 4-1 Parámetros de tecnologías de generación modelo PGA de prueba.

Se considera un perfil de demanda (curva azul de la Ilustración 4-1) cuyo promedio asciende a 13 GW con un factor de carga 0.83. Mientras que el perfil solar (curva roja perfil rojo en Ilustración 4-1) obtenido del explorador solar del departamento de geofísica Universidad de Chile, y se pondera por un factor para obtener un factor de planta de 0.4 [32].

Por otro lado, las curvas de viento son el resultado de aplicar un modelo ARMA [31] a series de viento, cuyas constantes son fijadas de modo de representar un parque eólico en Nueva Zelanda. Y algunos de sus ejemplos se observan en la Ilustración 4-2.

Por su parte, los parámetros correspondientes a los sistemas de almacenamiento están dados por la (tabla 4-2).

Finalmente, se agrega la posibilidad de desprender carga a un costo de 380 [USD/MWh]. No se modela el sistema de transmisión, evitando los efectos de ésta en las conclusiones.

Parámetro	Central de bombeo	Sistemas de baterías.
<i>Costos por potencia [USD/kW]</i>	42	28.58
<i>Costos por energía [USD/kWh].</i>	0.064	57.1668
<i>Eficiencia global.</i>	0.8	0.9

Tabla 4-2 Parámetros de los sistemas de almacenamiento modelo PGA de prueba.

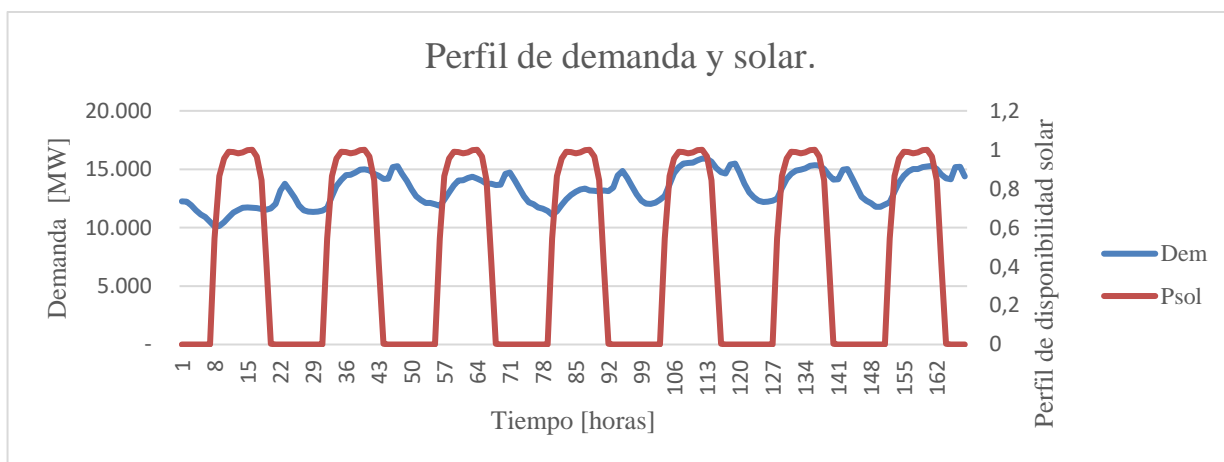


Ilustración 4-1 Perfil demanda y solar modelo PGA de prueba.

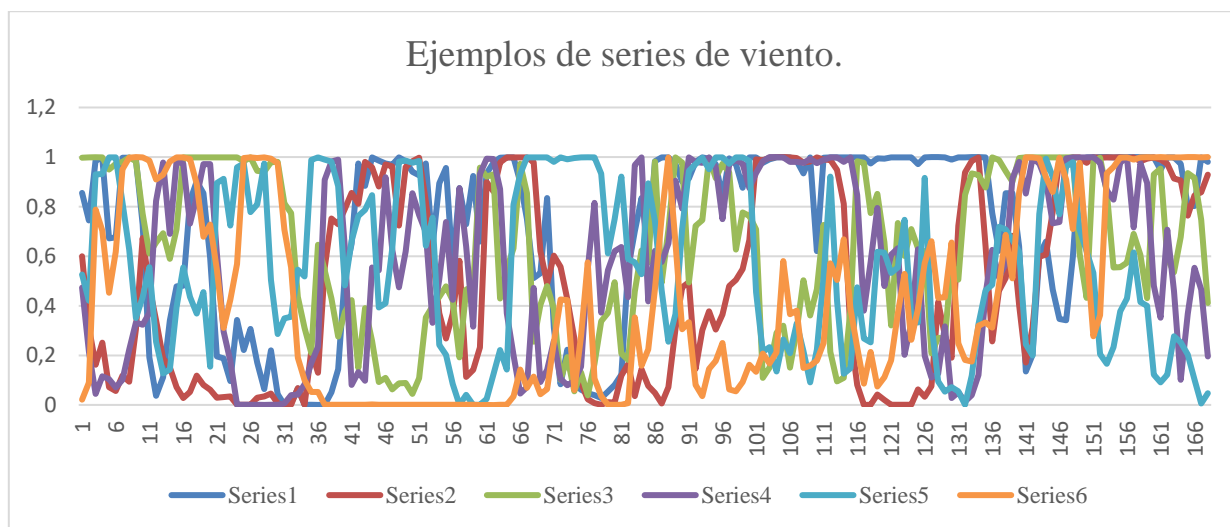


Ilustración 4-2 Ejemplos de serie de viento modelo PGA de prueba.

4.2. Modelo PGA aplicado al Sistema Eléctrico Nacional chileno

Este modelo tiene como objetivo verificar la potencialidad de la metodología propuesta en un problema de un tamaño considerable, y que sus conclusiones permitan ver el valor de la variable aleatoria en la configuración del mix óptimo que minimice el costo en el mercado eléctrico chileno.

Con este objetivo los datos que alimentan el modelo tienen sus siguientes fuentes mostradas en la Tabla 4-3.

Parámetro	Fuente
Costos de combustible	Comisión Nacional de Energía [33].
Costos de inversión	Centro de Cambio Global, Pontificia Universidad Católica [34].
Demanda	Proyección según informes del Coordinador Eléctrico Nacional.[35], [36]
Potenciales Energéticos	“El Potencial Eólico, Solar e Hidroeléctrico de Arica a Chiloé”, Ministerio de Energía y GIZ.[37]
Potencial Biomasa	Estudio de contribución de las ERNC al SIC al 2025, Universidad Técnica Federico Santa María. [38]
Costos de Encendido/ Apagado y rampas.	Centro de Energía, Universidad de Chile.

Tabla 4-3 Fuentes de los datos modelo PGA aplicado al SIN.

Los parámetros técnico-económicos que rigen a las centrales de carácter térmico se muestran en la Tabla 4-4. Cabe mencionar que el dato asociado a rampa horaria considera el tamaño de cada unidad, y de ahí que la rampa por unidad asociada a la tecnología con carbón tenga un mayor valor que la de las unidades a GNL

Por su parte, la Tabla 4-5 indica los parámetros técnicos de los sistemas de almacenamiento, que incluyen centrales de bombeo, sistemas de baterías y embalses que no tienen la posibilidad de cargarse directamente de la red.

Finalmente, la tabla 4-6 considera los parámetros de las tecnologías de FEV. Se considera la generación eólica, solar e hidráulica de pasada.

Tecnología	Carbón	GNL Ciclo Combinado	Diesel	Biomasa
<i>Costos de inversión [USD/kW]</i>	250	77.5	66.6	56.58
<i>Tamaño de cada unidad [MW]</i>	400	200	50	50
<i>Mínimo técnico de las unidades [MW]</i>	160	40	5	10

<i>Potencia máxima instalable. [MW]</i>	20.000	20.000	20.000	5.639
<i>Costos de operación [USD/MWh]</i>	43	171	230	18
<i>Costos de Encendido/ Apagado [USD]</i>	58.800	11.294	2.019	0
<i>Mínimo tiempo de encendido/Apagado</i>	24	2	0	0
<i>Rampa máxima horaria [MW]</i>	530	390	600	600

Tabla 4-4 Parámetros técnico económicos centrales térmicas del modelo PGA aplicado al SIN.

Tecnología	Sistema de baterías.	Centrales de Bombeo	Embalses
<i>Costos de inversión [USD/kW]</i>	52.5	143	216.6
<i>Tamaño de cada unidad [MW]</i>	100	500	500
<i>Tiempo de regulación [horas]</i>	2	168	672
<i>Potencia máxima instalable [MW]</i>	20.000	20.000	20.000
<i>Rampa horaria [MW]</i>	100	100	100
<i>Eficiencia de carga.</i>	0.98	0.9	-
<i>Eficiencia de descarga</i>	1.03	1.18	1.18

Tabla 4-5 Parámetros técnicos económicos sistemas de almacenamiento modelo PGA aplicado al SIN.

Tecnología	Eólica	Solar	Hidráulica de pasada.
<i>Costos de inversión [USD/kW]</i>	108.3	100	258.3
<i>Tamaño de cada unidad instalada [MW]</i>	100	150	500
<i>Potencia máxima instalable [MW]</i>	37.477	1.237.903	20.000

Tabla 4-6 Parámetros técnico económicos de tecnologías de FEV modelo PGA aplicado al SIN

La demanda corresponde a un vector de 4380 componentes (la cantidad de horas de seis meses) con un factor de carga 0.86. Mientras que la demanda máxima es de 16.552 MW. El recurso solar se obtiene a partir del explorador solar de la universidad de Chile [2]. Y tiene un factor de planta del 30%. Las centrales hidráulicas de pasada tienen una disponibilidad promedio de 32.59 MWh, es importante considerar que este valor subestima el potencial nacional. La Ilustración 4-3 da cuenta de las curvas equivalentes a una semana.

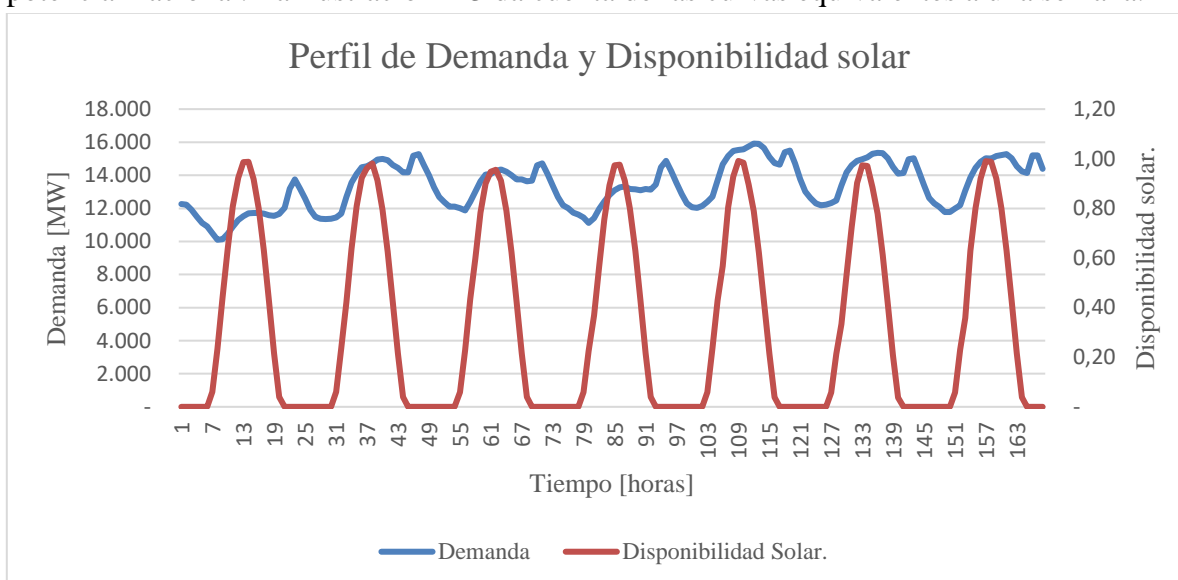


Ilustración 4-3 Perfil de demanda y disponibilidad solar de una semana modelo PGA aplicada al SIN.

Del mismo modo, con el perfil de viento tomado del explorador eólico de la Universidad de Chile [39], se generan 100 muestras de viento con el modelo ARMA, los cuales servirán de datos de entrada al modelo PGA. La ilustración 4-4 da cuenta de 6 curvas de una semana y muestran la variabilidad de las curvas.

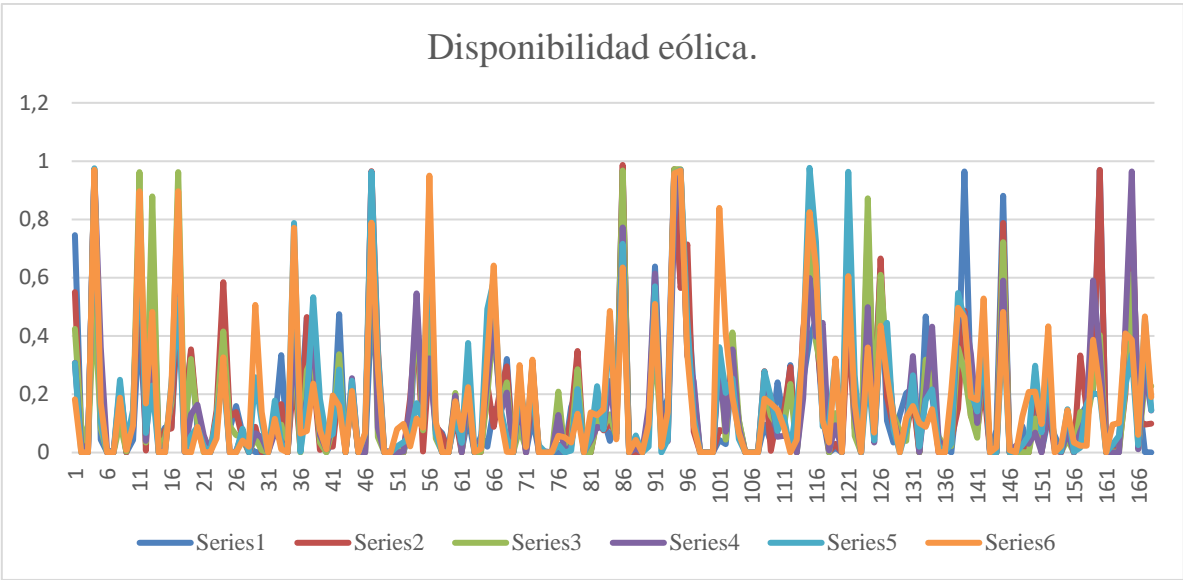


Ilustración 4-4 Disponibilidad Eólica modelo PGA aplicado al SIN.

5. Resultados y Análisis

Para analizar el funcionamiento de la metodología, se analizan dos casos de estudio denominados PGA de prueba y PGA aplicado al sistema eléctrico nacional (SEN) chileno.

Si bien ambos modelos aplican la metodología propuesta, el objetivo que tienen ambos son distintos. El modelo de prueba pretende mostrar el funcionamiento de cada fase de la metodología, junto con esto comparar los resultados con la programación estocástica hecha a partir de la totalidad de los escenarios.

Por su parte, el modelo PGA aplicado al SEN busca evaluar el potencial de la metodología aplicada a un modelo de planificación con un horizonte más extendido y acorde a un modelo de planificación con restricciones de corto plazo, junto con esto, se comparan los resultados con algoritmos existentes en la literatura, en particular el denominado *fast forward selection* propuestos en [16].

Finalmente, a modo resumen se hace una visión global del algoritmo de reducción de escenarios.

5.1. PGA de prueba

Tal como se indica en el capítulo 3, la metodología contempla el cómputo de versiones determinísticas y estocásticas de los modelos de planificación de la generación y almacenamiento con respecto a 100 distintas series de viento. Es importante considerar, que esta cantidad de viento se supone representativa de la variable aleatoria. También es importante considerar que la metodología de reducción de escenarios se justifica cuando la cantidad total de escenarios exige una carga computacional superior a los recursos que se tienen.

Si bien el PGA de prueba no tiene como objetivo analizar el mix, puesto que los parámetros de entrada como los costos no son validados, si permite analizar la problemática que la metodología busca resolver.

Al computar los escenarios determinísticos (en este caso se ejecutaron 100 simulaciones, una por cada serie de viento y se promediaron sus resultados) se encuentra que el mix óptimo promedio es pronunciadamente eólico, y existe una gran participación de centrales de bombeo. La ilustración 5-1 muestra el promedio de las soluciones del PGA.

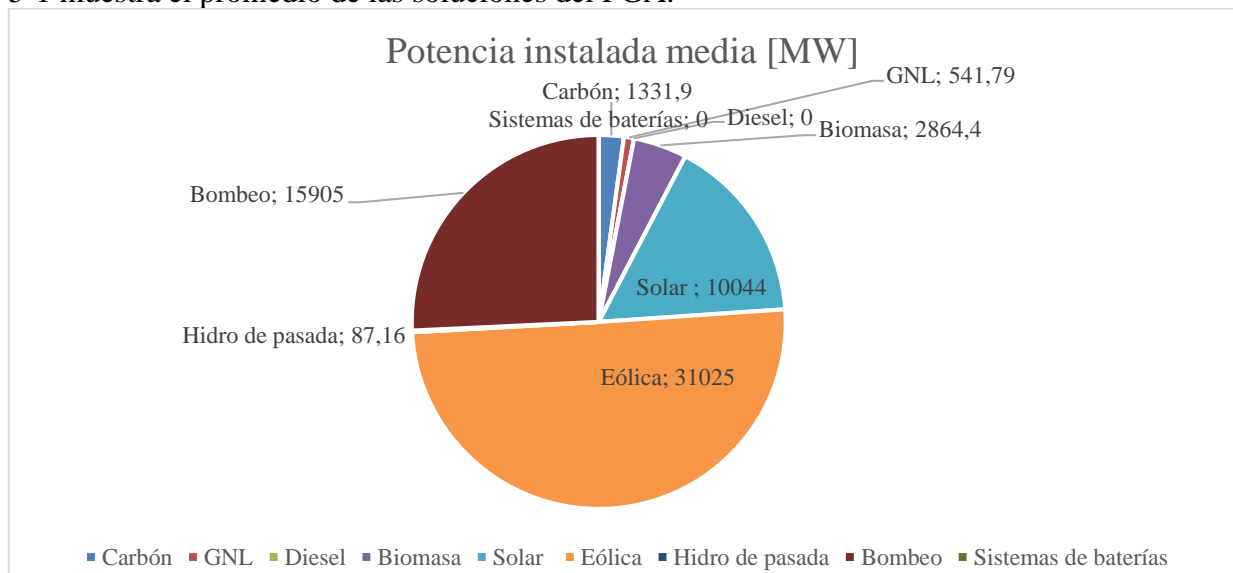


Ilustración 5-1 Mix promedio sistema de PGA determinísticos de prueba.

Sin embargo, al analizar su desviación estándar, es posible observar la variabilidad que tienen los escenarios determinísticos. Donde para el caso del viento alcanza alrededor del 25% y para el bombeo alcanza 18%. Esto permite concluir que en muchos casos las ejecuciones determinísticas inducen a discrepancias considerables en la conformación del parque generador.

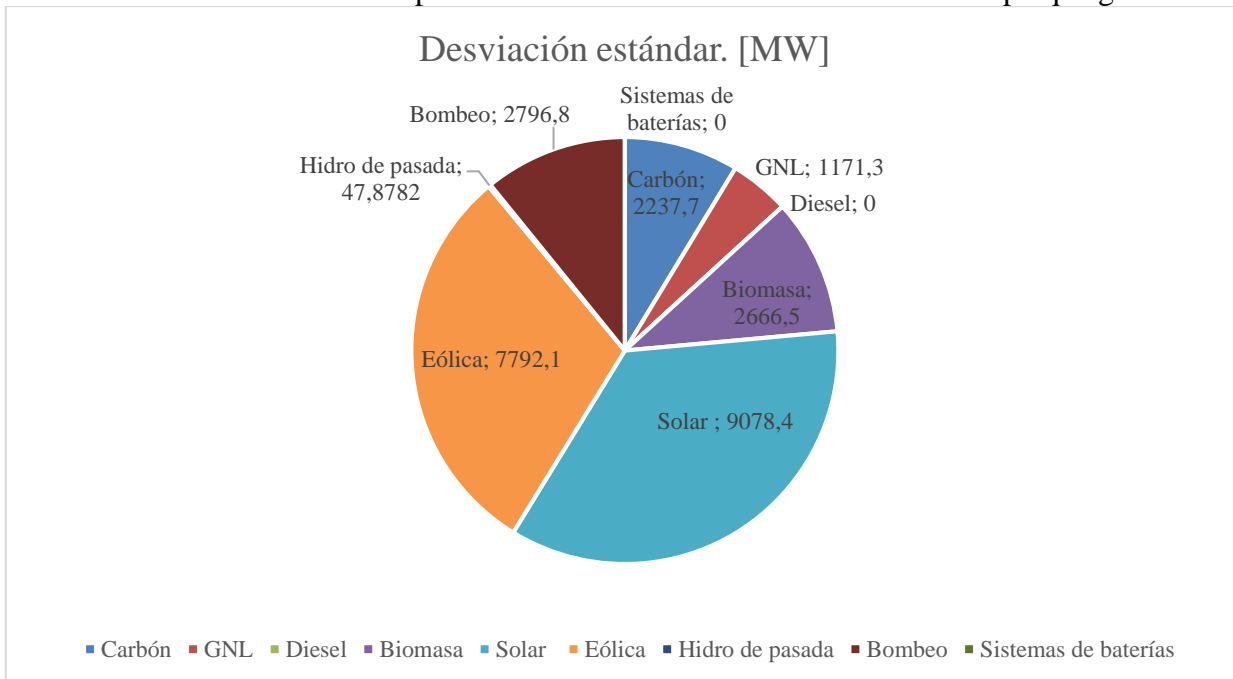


Ilustración 5-2 Desviación estándar asociada al mix óptimo aplicado al PGA de prueba.

En el PGA de prueba se tiene la posibilidad de ejecutar el problema completo con 100 escenarios, donde se llega al mix de generación-almacenamiento que se muestra en la Ilustración 5-3. Se observa en esta figura el parque generador alcanzado, cuyo costo total (costos de inversión más los costos de operación, ascienden a 11.23 MM USD, lo que representa un costo promedio de la energía de 4.95 USD/MWh. Es importante destacar que los precios no son representativos de la vida real, por lo que estos valores no tienen interpretación económica real.

Es posible observar en la Ilustración 5-3 el aumento de participación en el mercado de las tecnologías de GNL y una disminución en la participación del carbón. Del mismo modo, se observa una mayor capacidad instalada que permite abastecer energía en todos los escenarios. De este modo es posible observar además, que la diferencia entre las ejecuciones determinísticas y estocásticas, es que en las primeras se tiende a subestimar la participación de tecnologías que en ambientes de incertidumbre ofrecen al sistema una flexibilidad que permita trabajar al sistema en promedio a mínimo costo. En este caso se resolvió el problema estocástico sin una reducción de escenarios.

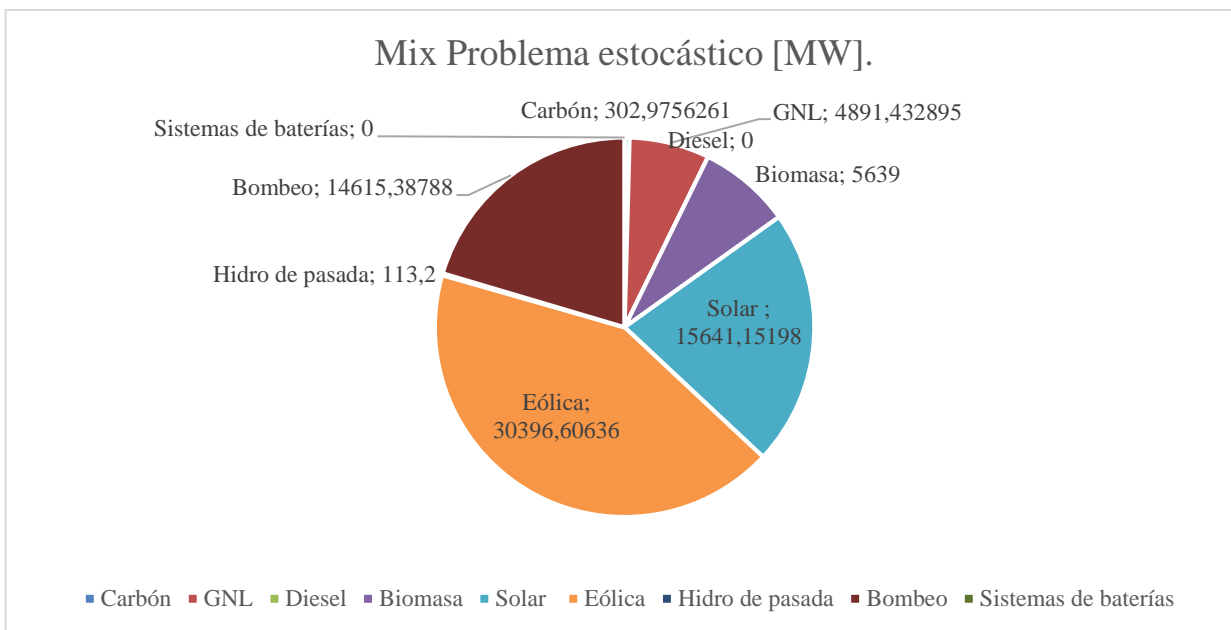


Ilustración 5-3 Mix óptimo del PGA estocástico aplicado al caso de prueba con 100 escenarios.

Es entonces que se hace pertinente una modelación estocástica con uso eficiente de los escenarios debido al ya comentado compromiso entre la exactitud y precisión de los resultados y la carga computacional necesaria para el cómputo de los cálculos.

5.1.1. Fase I: Selección de características

Se propone como “características” a analizar la capacidad instalada en cada una de las tecnologías y los costos totales de planificación en cada uno de los escenarios. La información mutua entre dichas características y las series de viento se puede observar en la tabla 5-11.

Tecnología	Información mutua.
Carbón	0.35
GNL	0.35
Diesel	0.00
Biomasa	0.20
Solar	0.53
Eólica	0.62
Hidráulica de pasada	0.54
Bombeo	0.82
Sistemas de baterías	0.00
Costos totales	0.93

Tabla 5-11 Información mutua entre las series de viento y características de PGA de prueba.

Es posible ver que las tecnologías que presentan menor información mutua con el viento son la biomasa, Diésel y los sistemas de baterías. En el caso de la primera, en la mayoría de los escenarios determinísticos se decide instalar todo el potencial de la tecnología, mientras que para las otras dos, se decide no invertir en dichas centrales. Tanto la biomasa, como el Diésel son tecnologías extremas bajo el punto de vista de una lista de mérito, mientras la primera actúa como

central de base, la segunda actúa en punta. La capacidad instalada en estas tecnologías no se ven tan afectadas como se ven las centrales que son medias en la lógica de la lista de mérito, como lo sería el GNL o el carbón (dicha tecnología está por sobre la biomasa).

Las tecnologías a base de FEV reciben una explicación similar en relación al valor de su información mutua, centrales con costos de inversión y disponibilidades similares, reciben un valor de información similar.

Finalmente, dado que no se modelan límites de reserva por tecnología, los sistemas BESS no resultan atractivos frente al bombeo, puesto que los primeros requieren una inversión considerablemente alta comparada con el bombeo y en una PGA sin restricciones de largo plazo, no ofrecen ventajas comparativas. De este modo, los excedentes del viento son tomados directamente por el bombeo, de modo que, a mayor disponibilidad, mayor inversión en viento, mayor excedente y luego más atractivo es la inversión en bombeo. Lo anterior no se da en todos los casos, sin embargo, es una intuición que se acompaña por las densidades de probabilidad de la ilustración 5-4.

La ilustración 5-4 da cuenta de las distribuciones de probabilidad con las cuales se calcula la información mutua de la tabla 5-1 Dichas distribuciones representan la capacidad instalada en cada una de las tecnologías normalizada por su media y desviación estándar. Se puede observar que las características que presentan un mejor comportamiento respecto a la información mutua son aquellas que tienden a tener más que dos *peaks*, y logran capturar de mejor manera la información contenida en las series de viento.

Resulta destacable entonces que la información netamente estadística tiene un sentido económico, por tanto, este hecho le da cierta racionalidad al índice propuesto en la ecuación 3-14 con el denominado índice estadístico ΔS_3 que combina la información estadística y económica.

Sin perjuicio de lo anterior, la mejor característica para este caso resulta ser la asociada a los costos totales de planificación, debido a que presentan la mayor información mutua; 0.93. Esta captura de mejor manera la variabilidad del recurso eólico, de modo que es esta la que se selecciona para la Fase II.

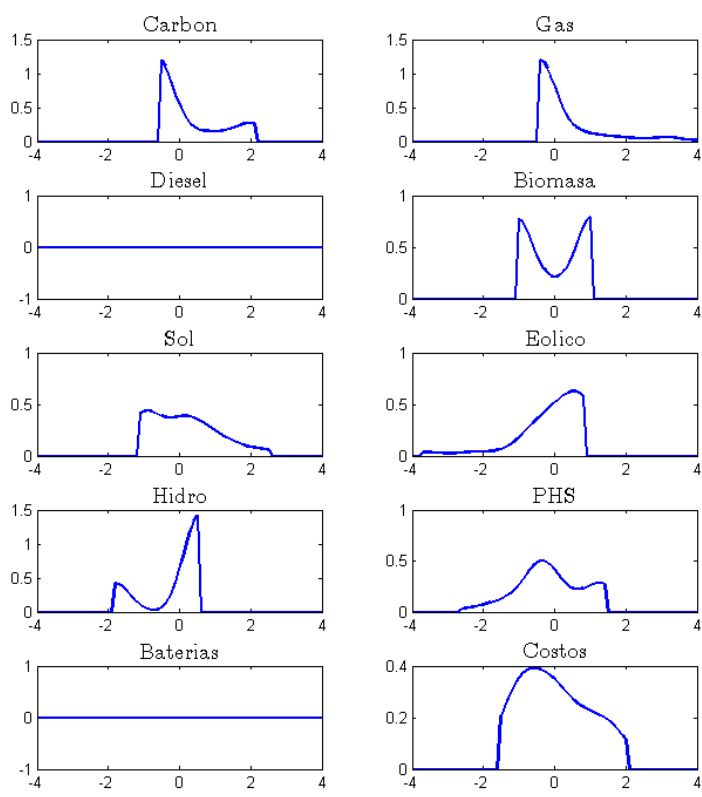


Ilustración 5-4 Densidad de probabilidad de las características normalizadas (PGA de prueba).

5.1.2. Fase II: Reducción de escenarios

Una vez seleccionada la característica corresponde la fase de Reducción de Escenarios. Con el objetivo de verificar los múltiples mínimos locales que puede llegar a tener la divergencia de Cauchy-Schwarz se grafica la función divergencia para el caso en que se realiza una reducción a 3 escenarios, lo que se puede observar en la ilustración 5-5. Donde es posible observar al menos 2 mínimos locales, e incluso puntos silla donde el método del gradiente descendente sin ningún tipo de ruido AWGN no evolucionaría en caso de caer en dicho punto.

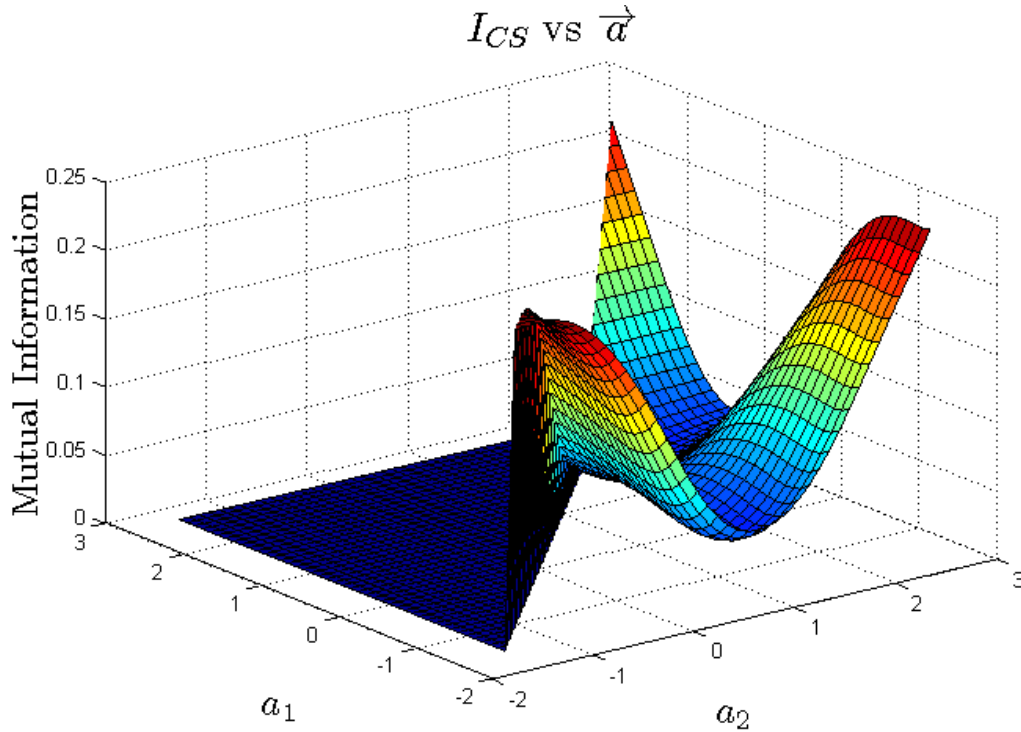


Ilustración 5-5 Divergencia de Cauchy Schwarz para el caso de reducción a 3 escenarios (PGA de prueba)

Más aun, dicha ilustración permite evaluar el buen funcionamiento de dicho funcional para conseguir distribuciones cercanas en el espacio de probabilidad. Es conveniente recordar que los puntos de \vec{a} son los que definen donde se inicia y termina cada *cluster*, si estos puntos se ubican a los extremos del dominio de la densidad de probabilidad entonces la distancia entre la distribución de los *cluster* crece, mientras que, los puntos \vec{a} ubicados en zonas intermedias suelen ser más representativos, y se refleja en la información mutua que tiene valores menores.

Para este problema se consideran 1000 iteraciones del gradiente descendente propuesto en la metodología (Sección 3.2), y el ruido blanco aditivo gaussiano se aplica con una relación señal a ruido (SNR) de 30dB. La ilustración 5-6 muestra el proceso de convergencia del proceso donde, en general, se distinguen dos zonas:

- La primera de ellas es donde domina el ruido, y dice relación de aquellas iteraciones donde a_i^t explora regiones cercanas a mínimos locales. En estas zonas se puede observar que las actualizaciones son dominadas por el ruido (gráfico de $\|a^{t+1} - a^t\|$), donde se observa esencialmente ruido entre las actualizaciones. Mientras que la norma del gradiente (gráfico de $\nabla I_{CS}(a^t)$) muestra que el gradiente es de baja norma, lo que es coincidente con el hecho de que se está explorando un mínimo local.
- La segunda de las zonas, corresponde cuando las iteraciones hacen que los a_i^t exploren zonas lejanas a mínimos locales, en cuyos casos, la convergencia se ve afectada en mayor medida por la norma del gradiente $\nabla I_{CS}(a^t)$, la cual crece y tiene como efecto que el valor de $I_{CS}(a^t)$ converja nuevamente a una zona de mínimos locales.

Es entonces que la metodología usada para el gradiente descendente con un ruido en su convergencia otorga la posibilidad de que el algoritmo pueda converger a varios mínimos locales durante una cantidad fija de iteraciones. Sin embargo, no asegura la convergencia al mínimo global.

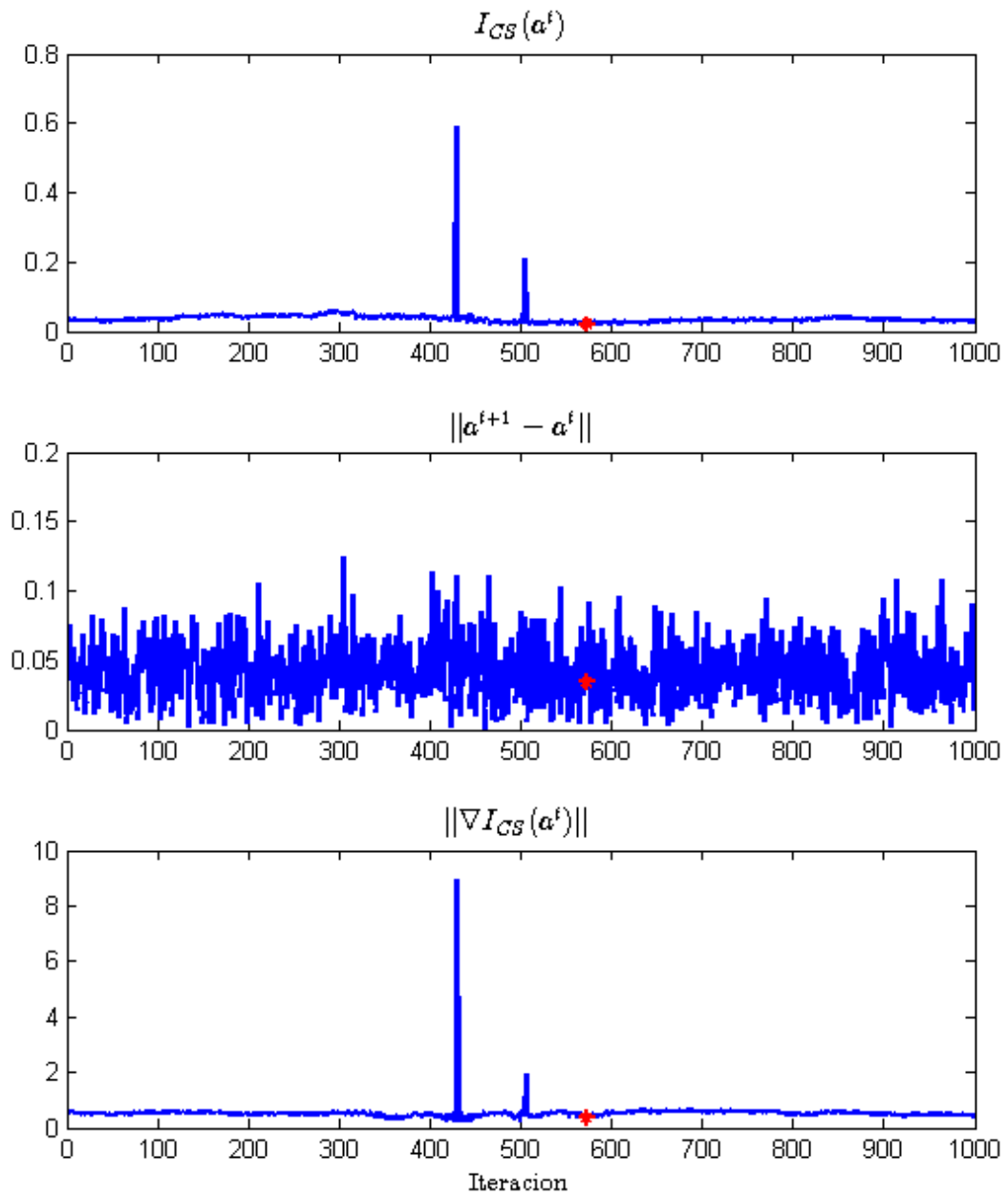


Ilustración 5-6 Convergencia del gradiente descendente para el caso PGA de prueba.

Aun cuando no se asegure la convergencia al mínimo global de I_{CS} , en muchos casos la solución es aceptable como se puede observar en la Ilustración 5-7 que muestra un ejemplo de la convergencia del algoritmo y la semejanza de esta solución con la solución óptima.

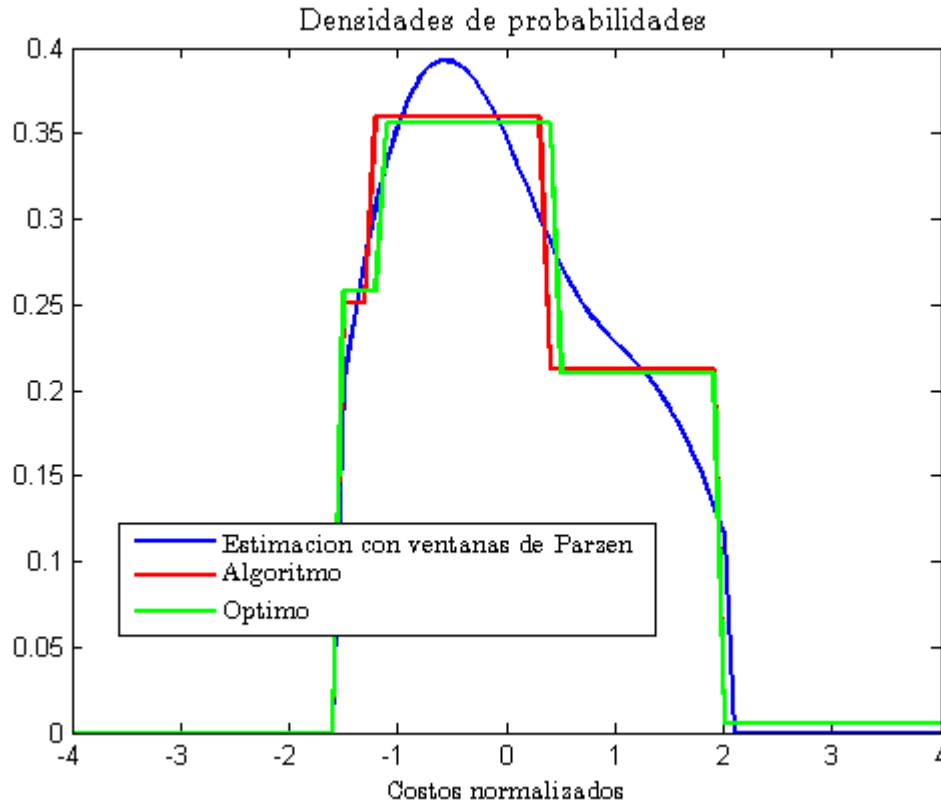


Ilustración 5-7 Comparación de densidad de probabilidad y su reducción óptima con la encontrada con el algoritmo.

Este análisis cualitativo de las soluciones se hace para el caso en que se tiene una reducción a 3 escenarios, y por lo tanto se requiere escoger dos puntos para distinguir los intervalos. El problema escala en dimensión según la cantidad de escenarios reducidos se quiere preservar, y por tanto no resulta efectivo encontrar la solución óptima exhaustivamente, en consecuencia, la comparación de los resultados no es posible llevarla a cabo.

Entre las posibles consideraciones de este algoritmo es que, a mayor cantidad de iteraciones, es más probable que el algoritmo alguna vez tome un valor cercano al óptimo gracias al movimiento aleatorio del ruido. De modo que el número de iteraciones está relacionado con el SNR del ruido Gaussiano (Mayor SNR implica una menor relevancia del gradiente, y por tanto la exploración es una mayor cantidad de iteraciones) y con la cantidad de escenarios que se preservan en el esquema de reducción (Mayor número de escenarios preservados, mayor cantidad de iteraciones).

5.1.3. Fase III: Análisis de post optimalidad y Fase IV: Ajustes

Para evaluar la precisión de las soluciones, los índices de precisión R_1 y R_2 son calculados para los casos en que la reducción se hace a $m = 3$, donde se puede observar que el índice R_1 que representa la diferencia relativa entre la cota inferior y superior de los costos totales de planificación representa un 12.5%, mientras que R_2 presenta un valor de 32% que indica el tamaño relativo de la derivada comparado con los costos de planificación.

En general, esta falta de precisión en los costos no es una variación aceptable lo que sugiere la necesidad de hacer ajustes que en este caso consideran el aumento de la cantidad de escenarios

reducidos a $m = 9$. Para esto nuevamente es ejecutada la fase 2. Se aprecia un rápido decrecimiento de los índices ($R_1 = 3.09\%$ y $R_2 = 9.63\%$). Análogamente, se procede a un nuevo aumento a $m = 27$ escenarios, sin embargo, el decrecimiento en este caso disminuye alcanzando valores ($R_1 = 1.2\%$ y $R_2 = -0.26\%$). Para esta cantidad de escenarios, en general se considera que la variación que señalan los indicadores está en rangos aceptables y se considera que el algoritmo ha convergido.

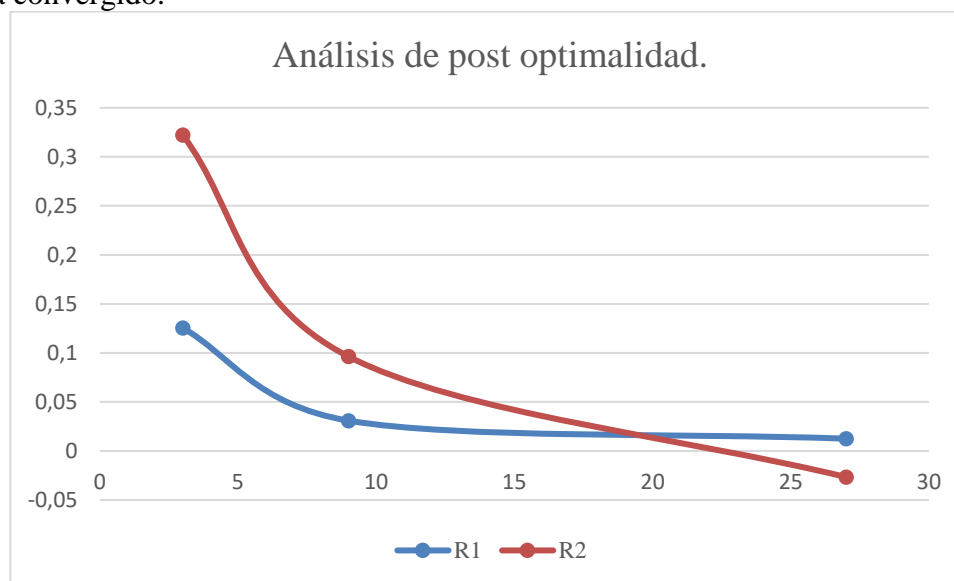


Ilustración 5-8 Análisis de post optimalidad PGA de prueba.

Del mismo modo, este análisis se extiende a los índices de adaptabilidad para las tecnologías de almacenamiento propuestos en las ecuaciones 3-12 a 3-14, los cuales otorgan una referencia de cuánto podría llegar a cambiar la capacidad instalada en sistemas de almacenamiento con bombeo. Los resultados se resumen en la Tabla 5-2. Donde se puede observar tendencias similares a los índices R_1 y R_2 en cuanto al fuerte decaimiento inicial y posterior decaimiento más suave. La tasa de decrecimiento inicial para R_1 es del orden de 0.12 por escenario en el primer tramo, mientras que en el segundo tramo la tasa corresponde a 0.03 por escenario. De manera análoga, para R_2 en el primer tramo, la tasa de decaimiento corresponde a 0.117 por escenario, mientras que, en el segundo tramo de la curva, la tasa de decaimiento corresponde a 0.068 por escenario

m	ΔS_1 [MW]	ΔS_2 [MW]	ΔS_3 [MW]
3	10087	4218	3102
9	4160	579	604
27	1687	339	300

Tabla 5-2 Índices de adaptabilidad para el PGA de prueba.

En este caso, ΔS_1 representa una estimación estilo “peor caso” respecto a la variación de almacenamiento al agregar un escenario. Por cierto, es el indicador con mayor valor de los tres indicadores propuestos. ΔS_2 y ΔS_3 , aun cuando son indicadores de naturaleza distinta, presentan valores similares lo que permite intuir que los valores económicos (costo marginal de la tecnología) con los cuales se calcula ΔS_2 no son ajenos al cálculo del indicador estadístico ΔS_3 . La razón es que las tecnologías en La ilustración 5-4 da cuenta de las distribuciones de probabilidad con las cuales se calcula la información mutua de la tabla 5-1 Dichas distribuciones representan la capacidad instalada en cada una de las tecnologías normalizada por su media y desviación estándar.

Se puede observar que las características que presentan un mejor comportamiento respecto a la información mutua son aquellas que tienden a tener más que dos *peaks*, y logran capturar de mejor manera la información contenida en las series de viento.

Las que más invierte en el sistema, son aquellas que económicamente reportan más beneficio (costos marginales más bajos) y coincide con aquellas que presentan mayor información mutua al hacer el cálculo de I_M en la Tabla 5-1.

5.1.4. Revisión general del PGA de prueba.

Finalmente, dado que los indicadores con $m = 27$ presentan valores bajos, se puede considerar que con esta cantidad de escenarios el algoritmo ha convergido. En este caso es posible comparar los resultados con el árbol determinístico equivalente con 100 escenarios. Los resultados se pueden observar en la Tabla 5-3.

m	$ \hat{s} - s^* $ [MW]	Costos totales [MM USD]	Error de costos relativo
27	361	11.0	1.81%

Tabla 5-3 Comparación solución de reducción de escenarios y cantidad total de escenarios.

La diferencia de costos es del orden de R_1 (cuyo valor es 1.25%) es un buen estimador del error de costos relativos. Es importante notar que R_1 representa el máximo error relativo cuando se aumenta en uno la cantidad de escenarios. Sin embargo, en la práctica, demuestra ser un buen indicador para un mayor salto de pasos.

La diferencia de potencia instalada es menor a ΔS_1 (como era de esperar), y es del orden de ΔS_2 y ΔS_3 . Dicha diferencia representa el 2.4% de la potencia instalada (Ver Ilustración 5-3). Esto nos permite analizar la bondad de los estimadores, y la cercanía de estos con el error de estimación que se comete al reducir los escenarios.

Conforme al error relativo disminuye, también disminuye el valor de los indicadores de adaptabilidad ΔS_i y las estimaciones de la capacidad instalada son cada vez más precisas. Sin embargo, la cota para cada uno de los indicadores debe ser de acuerdo al indicador en particular.

5.2. PGA aplicado al SEN.

A diferencia del modelo anterior, este modelo cuenta con costos y datos representativos para el sistema eléctrico nacional chileno. Sobre este modelo el desafío de ejecutar el PGA con 100 escenarios en un computador de escritorio no es factible, de modo que el uso de técnicas de reducción de escenarios es imperioso.

Es en este modelo entonces que se realiza la comparación con el algoritmo *Forward Selection* propuesto en [16] sobre los índices R_1 y R_2 y los índices de adaptabilidad aplicados a las tecnologías de almacenamiento aplicados a las centrales de bombeo hidráulico y los sistemas de baterías s

En lo sucesivo se analizan los resultados para las fases de la metodología y finalmente se hace una comparación con el mencionado algoritmo de la literatura.

5.2.1. Aplicación de la metodología en el modelo del SEN chileno.

Los resultados asociados a la Fase I de Selección de Características se resumen en la Tabla 5-4.

Característica	Información Mutua.
Carbón	0.00
GNL	0.24
Diesel	0.06
Biomasa	0.00
Solar	0.78
Eólica	0.95
Hidráulica de pasada	0.00
Bombeo	0.002
Sistema de baterías	0.58
Embalse	0.00
Costos totales	0.58

Tabla 5-4 Información mutua entre las series de viento y características del SEN chileno.

De manera similar al PGA de prueba, las tecnologías extremas en la orden de mérito (Tecnología de base como es el carbón, o la tecnología de punta como puede ser el Diésel) son débilmente afectadas por la disponibilidad del recurso eólico. Lo anterior puede encontrar su explicación al analizar el parque que se obtiene y su despacho (ilustraciones 5-9 y 5-10) donde se puede observar que la inversión en centrales eólicas asciende a 100 MW, en un parque que está constituido por 19550 MW. Esta baja inversión se puede explicar por la serie de viento tomada del Explorador eólico [39] (En la ubicación del actual parque de canela. ver ilustración 4-4) corresponde a un viento con una variabilidad muy grande, que dificulta su incorporación en el SEN debido a la alta flexibilidad que el sistema requeriría. De modo que, a estos niveles de penetración eólica, el despacho de las centrales carboneras y de biomasa no se ve afectado en mayor medida por la disponibilidad del viento, y, por tanto, la inversión tampoco. A diferencia de la tecnología GNL, la cual capturaría las variaciones del viento. (Observar como en la hora 48 de la Ilustración 5-10 el GNL recorta puntas).

Por otro lado, la tecnología solar fotovoltaica aparece como una tecnología sustituta a la tecnología eólica, de modo que aumentos en inversión en una se obtienen a cambio de un desmedro en la inversión en la otra tecnología. Su explicación de la alta información mutua es más bien económica, debido a que mejores vientos, hacen las centrales eólicas más competitivas respecto a las solares, de modo que su penetración en el sistema aumenta.

Finalmente, debido a la alta variabilidad de las series de viento, se produce un alto complemento entre los sistemas de baterías y las centrales eólicas. Esta conclusión se obtiene debido a que el servicio que tienden a ofrecer los sistemas de baterías está asociado a los servicios complementarios (Observar hora 48 de la Ilustración 5-10, en que de manera análoga al GNL, el sistema de baterías recorta puntas). De este modo, mayor penetración eólica demanda una mayor demanda por servicios complementarios, y luego las baterías se hacen más competitivas.

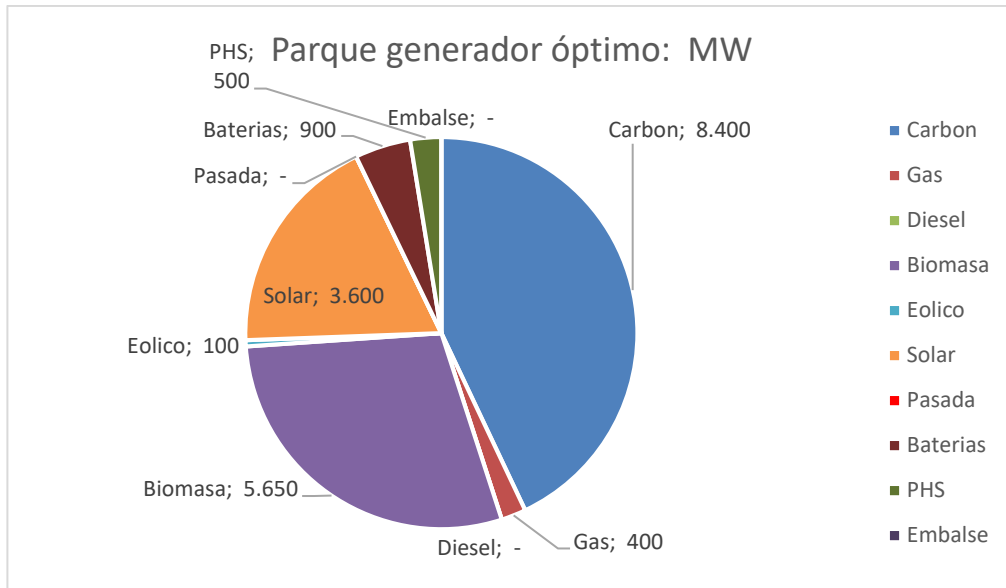


Ilustración 5-9 Mix óptimo obtenido con 3 escenarios aplicado al SEN chileno.

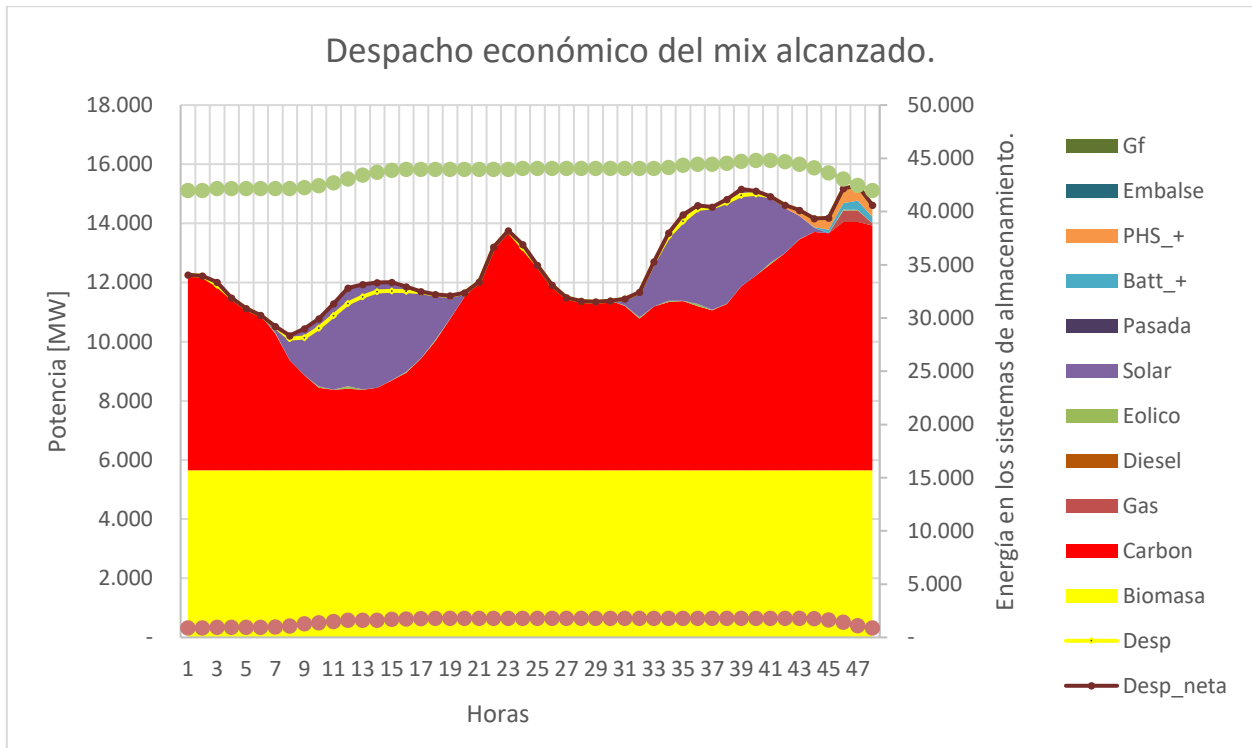


Ilustración 5-10 Despacho económico de dos semanas del mix óptimo alcanzado.

Para este caso, la característica que mejor recoge la información que contiene el viento es la capacidad instalada en viento. Si bien este resultado no siempre es así (de hecho, un ejemplo es el modelo de prueba), resulta intuitivo pensar que la capacidad que se invierte en centrales eólicas está fuertemente influenciada por la disponibilidad del viento.

Con esto se puede aplicar la Fase II de reducción de escenarios con una reducción a un $m = 3$ escenarios, y posteriormente, en la Fase III se ejecuta el Análisis Post optimal. Al calcular los índices de precisión se obtiene $R_1 = 0.0012$ y $R_2 = 0.0046$. Dichos valores se pueden considerar

como que el algoritmo ha convergido. Posteriormente, se puede calcular los índices de adaptabilidad ΔS_1 , ΔS_2 y ΔS_3 para los sistemas de baterías y la tecnología de bombeo. Los resultados se resumen en la Tabla 5-5 donde además se hace la comparación con el algoritmo *forward selection* (FS).

Tecnología	ΔS_1 [MW]	ΔS_2 [MW]	ΔS_3 [MW]
Sistema de baterías	3538	30	-41.42
Bombeo	561	17	0
Sistema de baterías/ FS	12619	30.81	-37.97
Bombeo/ FS	603.21	17.1	0

Tabla 5-5 Índices de adaptabilidad para la reducción de escenarios a $m=3$.

Cabe resaltar que si bien ΔS_1 presenta altos valores, para el caso de la tecnología bombeo la variación es del orden de la capacidad instalada. Dada la interpretación de este índice, que indica un “peor escenario” respecto a la variación de inversión, los valores encontrados son valores aceptables.

Por su parte, ΔS_2 entrega valores en torno al 3% de la capacidad instalada para el caso del sistema de baterías y el bombeo. En muchos casos, este valor es suficientemente pequeño con el cual se puede entender que la metodología ha convergido.

Finalmente, ΔS_3 que representa el índice estadístico entrega una variación de 4.6% para el sistema de baterías. Mientras que para el sistema de bombeo no reporta variaciones, puesto que en la Fase I se encuentra que la información mutua entre las series de viento y la capacidad instalada en dicha tecnología es cero, por lo que se estima que son variables independientes y se considera que la capacidad instalada corresponde a 500 MW.

Respecto a los tiempos computacionales, es importante mencionar que los experimentos de la Fase I son programables mediante computación paralela., de todos modos, para el caso chileno toman en promedio 30 segundos la resolución de cada uno de los PGA determinísticos. La fase de Selección de escenarios es de rápida resolución, menor a los 5 segundos, aunque está sujeta al número de iteraciones que se aplica al gradiente descendente. La etapa más crítica corresponde al cálculo de indicadores de adaptabilidad, puesto que requiere la resolución de un problema estocástico. Con 3 escenarios toma un tiempo cercano a los 1000 segundos, con 4 el tiempo sube a 3000 segundos mientras que con 7 escenarios el tiempo escala a los 20000 segundos. En esos contextos, incluso se requiere una alta capacidad de memoria RAM para tener capacidad de cálculo, para este problema se usó alrededor de 58 GB con 7 escenarios. No se tuvo la capacidad de cálculo para una mayor cantidad de escenarios. Los problemas GSEP son programados en el software FICO, mientras que el resto de la metodología es implementada en Matlab.

Lo anteriormente radica en la formulación de las restricciones de corto plazo obliga la resolución de una optimización del tipo lineal entera mixta (MILP por su acrónimo en inglés) lo cual es necesario en sistemas eléctricos como el sistema chileno, en donde existen unidades con gran capacidad y su predespacho resulta relevante en la operación. Mas un MILP no es necesariamente necesario en sistemas grandes interconectados, en donde el encendido o apagado de una unidad del sistema no afecta de mayor manera la operación. [40]

La comparación con el algoritmo *fast forward selection* (FFS) puede ser a partir de muchos puntos de vista. Respecto a los tiempos computacionales, los algoritmos propuestos en la literatura son más rápidos. Sin embargo, esta comparación no considera el lazo que ajusta la cantidad de escenarios dado un nivel de precisión y exactitud requeridos, de modo que FFS puede llegar a resultados con poca representatividad de los escenarios en el caso de partir con un m pequeño, o sobre representatividad en caso contrario, teniendo un PGA que toma más del tiempo necesario en

ser resuelto. En esa línea el algoritmo *FFS* es comparable con la Fase II de reducción de escenarios, puesto que ambos reciben como entrada escenarios (en esta tesis representados por características) y entregan los escenarios más representativos de la muestra. Resulta ser que tanto *FFS* como la Fase II toman poco tiempo, en el orden de segundos. Del mismo modo, al comparar el algoritmo con *fast forward selection* con respecto a los indicadores de adaptabilidad, la metodología ofrece mejoras sustanciales con respecto a el índice de adaptabilidad ΔS_1 , mientras que, en el resto de los índices, no se observan mayores mejoras. En tanto para los índices de precisión se obtienen los mismos resultados obtenidos con la metodología propuesta.

Para verificar la efectiva convergencia de los resultados se analiza una reducción a 4, 5 y 7 escenarios. Observándose los índices de la tabla 5-6.

Índice	$m = 4$	$m = 5$	$m = 7$
$\Delta S_1 BESS$	2460	2050	1540
$\Delta S_1 PHS$	500	420	310
$\Delta S_2 BESS$	30	25	20
$\Delta S_2 PHS$	13	14	10
$\Delta S_3 BESS$	0,7	0,6	0,4
$\Delta S_3 PHS$	0	0	0

Tabla 5-6 Índices de adaptación para el sistema eléctrico nacional chileno.

Cabe mencionar que la instalación óptima encontrada con escenarios corresponde a 900 MW de en sistemas de baterías (BESS por su acrónimo en inglés) y 500 MW de bombeo (PHS). En ese contexto, se observa que si el criterio de detención hubiese sido el indicador ΔS_1 , entonces más iteraciones hubiesen sido necesarias, puesto que el peor para aquel número de iteraciones es muy crítico, hecho que se acentúa cuando hay pocos escenarios, y agregar un escenario para el cálculo de los indicadores puede implicar que la muestra sea no representativa. Sin embargo, la ganancia en exactitud al agregar escenarios es ínfima comparada con la capacidad instalada cuando el criterio de detención lo imponen los indicadores 2 y 3, puesto que estos indicadores no son tan exigentes como el uno.

6. Conclusiones y Trabajo Futuro.

6.1. Conclusiones sobre la metodología.

Dada la importancia de incluir las restricciones operacionales de corto plazo en un PGA con alta penetración de FEV, la carga computacional se convierte en un desafío incluso para planificaciones determinísticas. Sin embargo, la naturaleza intrínseca de las FEV, que tiene incertidumbres asociadas, hace que trabajar con modelos estocásticos sea atractivo desde el punto de vista de una adecuada representación de dichas fuentes.

En ese contexto, se reconocen dos clases de reducción de escenarios en la literatura. La primera y más común es aquella que busca que los m escenarios reducidos sean cercanos a los n escenarios originales, sin tener en cuenta la optimización a la cual van a alimentar. Mientras que la segunda analiza cómo afecta cada escenario al problema de optimización, tomando distintos atributos de ejecuciones determinísticas del problema de optimización, para finalmente realizar un *clustering* con *k-means* bajo el supuesto inherente de que los *clusters* distribuyen como funciones Gaussianas (supuesto que en la práctica rara vez se cumple), y aunque no busquen tener cercanía en el espacio de probabilidad, indirectamente se logra con el uso de la regla de mínimo transporte o versiones análogas de ésta.

Esta tesis recoge elementos de ambos mundos, en el sentido que busca que los m escenarios reducidos tengan una cercanía en el espacio de probabilidad. Sin embargo, usa el contexto que otorga el problema de PGA. Es así que se realiza un *clustering*, pero sin caer en el supuesto de distribuciones Gaussianas, lo que le otorga una mayor generalidad a la metodología.

La propuesta en esta tesis se centra en una manera de tratar con escenarios de un modo eficiente, definiendo la menor cantidad de escenarios necesarios para satisfacer nociones de precisión y exactitud. Más aun, en muchos casos donde la ejecución de un modelo estocástico con un gran número de escenarios es no factible debido a la capacidad computacional disponible, el marco metodológico propuesto no solo encuentra una solución eficiente del mix óptimo (en términos de la calidad de la solución vs esfuerzos computacionales), sino que además propone una manera de encontrar un rango en el cual la capacidad instalada en almacenamiento podría cambiar si un escenario es añadido.

Aun cuando esta metodología es general y es válida para otras aplicaciones, el marco propuesto y sus casos de estudio (Atributos de la planificación y estimadores para éstos) son novedosos para la definición de un parque generador con almacenamiento óptimo, en el sentido de proponer índices de adaptabilidad de dichas tecnologías.

Más específicamente, para los casos de estudio que consideran series de disponibilidad de generación eólica en un período de 168 horas (una semana) y 4380 horas (seis meses), se encuentra que la cantidad de escenarios reducidos que permite tener una variación menor a un 2% de los costos totales son 27 y 3 escenarios respectivamente. Dichos resultados son explicables por la variabilidad del recurso eólico. En una semana, el viento puede llegar a ser muy variable, teniendo semanas con mucho viento, y otras con muy poco, lo que hace que se requieran más escenarios para una representación adecuada. A lo largo de un semestre, por otro lado, las variaciones se compensan, y por tanto se necesita una menor cantidad de escenarios para tener una representación adecuada. Esto permite concluir que el algoritmo es capaz de recoger la incertidumbre asociadas

al problema de optimización y en base a esto decidir la cantidad de escenarios necesarios para llegar a una representación adecuada de dicha variable aleatoria.

Sobre los indicadores de adaptabilidad de las tecnologías de almacenamiento, todos están asociados a una estimación de la cantidad que puede cambiar en la inversión del problema de optimización cuando un escenario es agregado. ΔS_1 suele presentar valores elevados, pues es una sobreestimación de la variación de inversión. Por otro lado, ΔS_2 y ΔS_3 empíricamente en los casos analizados, logran representar adecuadamente la variación de la inversión. Finalmente, dentro de cierto rango, se observa una semejanza de los últimos dos estimadores, incluso cuando el procedimiento en sus cálculos es distinto. ΔS_2 recoge aspectos económicos (valor de las variables duales) bajo la premisa de mantener la proporción de las capacidades de las distintas tecnologías. ΔS_3 por su parte recoge la información mutua de la capacidad instalada en una tecnología como medida de cuan sensible es dicha cantidad con respecto a la serie de viento. Esto último, si bien sigue una lógica estadística, los resultados muestran que no son ajenos a la racionalidad económica del proceso de planificación, puesto que la variación en inversión entre los distintos escenarios se hace respecto a aquellas tecnologías que marginalmente reportan mayores beneficios, y que resulta ser las más sensibles a las variaciones de la variable aleatoria. Sin embargo, también puede ocurrir que estos indicadores se desacoplen, en situaciones en que alguna tecnología no comparta información mutua con la variable aleatoria, o bien, que las señales económicas que se obtienen del PGA, no logren capturar el valor de agregar un escenario adicional.

Con el fin de comparar los resultados de la metodología propuesta, se compara ésta con la metodología *forward selection* con respecto al valor de los índices de adaptabilidad ΔS_i y con respecto al mix óptimo propuesto. En general se obtienen valores semejantes respecto a estos índices con excepción de ΔS_1 , donde la propuesta de la tesis evidencia mejoras sustanciales. Respecto al mix óptimo alcanzado, éste es esencialmente el mismo. Sin embargo, los algoritmos propuestos en la literatura en general no consideran un análisis post optimal y tampoco una estimación de variaciones en las capacidades instaladas que propone la metodología. En ese sentido, la Fase III y Fase IV se encargan de entregar robustez a los resultados obtenidos.

La Fase I y la Fase II cumplen el rol de algoritmos expuestos en la literatura como el *forward selection* o el *backward reduction*. Sin embargo, el uso de las metodologías propuestas puede llegar a un sobre requerimiento de escenarios al perder la noción de cómo afectan los escenarios al problema de optimización. Incluso la Fase II puede ser modificada por un *clustering* con *k-means*, sin embargo, esto no garantiza m escenarios reducidos cercanos en el espacio de probabilidad.

Esto refuerza el aporte de la metodología en conjunto, como una manera sistemática de encontrar la mínima cantidad de escenarios que satisfaga alguno de los índices de adaptabilidad propuestos, satisfaciendo nociones de precisión y exactitud.

6.2. Trabajo Futuro.

El trabajo futuro se centra en reconocer las debilidades de la propuesta y cómo mejorarlas. Y un segundo punto es analizar en que otros contextos dicha metodología puede ser aplicada.

Respecto a las debilidades se reconocen.

- Como mejora transversal al método se propone explotar el potencial de paralelización inherente a la metodología propuesta e incorporar el sistema de transmisión a los modelos de planificación

- Fase I: En este marco metodológico se proponen características como la capacidad instalada o los costos totales. La ventaja de este enfoque es que las características tienen su interpretación técnico-económica. Sin embargo, a cambio de perder esa interpretación, se puede analizar un funcional que comparta una mayor información mutua con la variable aleatoria. Con enfoques desde el mundo del aprendizaje de máquinas, es posible encontrar mejores características. Un ejemplo puede ser *Non Negative Matrix Factorization* para extraer características a partir de la salida del PGA. Como resultado de tener mejores características se espera que al hacer la reducción de escenarios es posible determinar *clusters* con una menor divergencia, por tanto, soluciones más exactas.
- Fase II: tal como se muestra en los resultados, con una baja cantidad de escenarios a reducir, el método del gradiente descendente con AWGN presenta una convergencia adecuada al mínimo global. Sin embargo, cuando se requiere un m mayor, el problema crece en dimensionalidad y su convergencia se hace más compleja. Dada la no linealidad de la divergencia, en la literatura se proponen métodos basados en algoritmos genéticos y enjambre de partículas, que pueden tener mejores desempeños en la minimización de dicho funcional.
- Fase III: Para el cálculo de los índices de adaptabilidad se hace uso de las variables duales asociadas a la inversión. La existencia de estas variables implica que el problema sea convexo, lo cual limita el uso de variables enteras necesarias en el *Unit commitment* de las restricciones de corto plazo. La forma de enfrentar dicha limitante en esta tesis consiste en ejecutar el PGA con una relajación lineal de las variables enteras. Esto puede inducir a valores alterados del cálculo de las variables duales. Si bien esta es una limitante teórica, se puede explorar distintas alternativas como realizar un cálculo por definición de la variable dual, esto es, ver cómo cambia la función objetivo cuando 1 MW adicional de la tecnología en cuestión es agregado.
- Fase IV: Como ya se analizó, las características propuestas pueden salir de algún algoritmo de “extracción de características” desde la Fase I. La otra vía de ajustes consiste en aumentar la cantidad de casos reducidos. Sin embargo, tal como se observa en la Ilustración 5-8, los criterios de detención pueden presentar tasas de decrecimiento alto, lo cual otorga un gran valor al escenario adicional, pero también se puede dar lo contrario. Esto implica que la cantidad de que se aumenta los escenarios puede tener un impacto variado en la reducción de los criterios de detención. Por lo que la noción de la “mínima” cantidad de escenarios se ve afectada si es que el aumento es demasiado grande. En caso contrario, un aumento pequeño de la cantidad reducida implicaría que el algoritmo converja en más iteraciones, lo cual implica un mayor gasto computacional.

En esta tesis la metodología propuesta es aplicada a problemas de PGA uninodal, sin embargo, se puede extender a una versión multimodal. Pero no solo a problemas de planificación, también es extensible a cualquier contexto en que la función que define el problema de optimización cumpla con la ecuación (2-64) que da cuenta de la Lipschitz continuidad. En ese sentido, resulta de interés evaluar el potencial de la propuesta en problemas con distinto horizonte de evaluación, como, por ejemplo, un problema de operación de una semana de un sistema eléctrico, o bien, un *day-ahead market* en donde las FEV podrían tener un mayor impacto y tener soluciones precisas de mayor interés. Esto puede tener un mayor

impacto en los costos del *real time market* al tener un mejor manejo de la incertidumbre que inducen este tipo de tecnologías.

Finalmente, en los casos de estudio de este trabajo se analiza como un PGA es sensible a distintos escenarios de viento. De modo que la variable que induce a un problema con incertidumbre es única. Sin embargo, en la práctica hay más fuentes de incertidumbre, como por ejemplo los costos de los combustibles, costos de inversión de algunas tecnologías etc. El algoritmo presentado es capaz de determinar cuál de todas las fuentes de aleatoriedad es más importante en el PGA (o problema de optimización a resolver) otorgando una mayor cantidad de escenarios en representar adecuadamente su aleatoriedad, mientras que variables menos relevantes reciben una menor cantidad de escenarios.

7. Bibliografía

- [1] D. Jitka, «Stability and sensibility analysis for stochastic programming,» *Annals of operations Research*, vol. 27, pp. 115-142, 1990.
- [2] Ministerio de Energía; Departament Geofísica, Universidad de Chile, «Explorador Solar,» [En línea]. Available: <http://ernc.dgf.uchile.cl:48080/inicio>. [Último acceso: 15 12 2017].
- [3] REN21, Renewables 2017 Global Status Report, Paris, 2017.
- [4] L. Xie, P. M. S. Carvalho, L. A. F. M. F. J. Liu, B. H. Krogh, N. Popli y M. D. Ilic, «Wind Integration in Power System: Operational Challenges and Possible Solutions,» *Proceedings of the IEEE*, vol. 99, n° 1, pp. 214-232, Jan. 2011.
- [5] North American Electric Reliability Corporation, «Accommodating High Levels of Variable Generation,» Princeton, NJ, April 2009.
- [6] R. Moreno, R. Moreira y G. Strbac, «A MILP model for optimising multi-service portfolios of distributed energy storage,» *Applied Energy*, n° 137, pp. 554-566, 2015.
- [7] J. Haas, F. Cebulla, K. Cao, W. Nowak, R. Palma-Behnke, C. Rahmann y P. Mancarella, «Challenges and trends of energy storage expansion planning for flexibility provision in low-carbon power systems - a review,» *Renewable an Sustainable Energy Reviews*, n° 80, pp. 603-619, 2017.
- [8] A. Flores-Quiroz, R. Palma-Behnke, G. Zakeri and R. Moreno, "A column generation approach for solving generation expansion planning problems with high renewable energy penetration," *Electric Power System Research*, no. 136, pp. 232-241, 2016.
- [9] B. Palmintier y M. Webster, «Impacto of Unit Commitment Constraints on Generation Expansion Planning with Renewables,» *2011 IEEE Power and Energy Society General Meeting*, 2011.
- [10] W. B. Powell y S. Meisel, «Tutorial on Stochastic Optimization in Energy -- Part I Modeling an Policies,» *IEEE Transactions on Power Systems*, vol. 31, n° 2, pp. 1459-1467, 2016.
- [11] W. B. Powell y S. Meisel, «Tutorial on Stochastic Optimization in Energy -- Part II: An Energy Storage Illustration,» *IEEE Transactions on Power System*, vol. 31, n° 2, pp. 1468-1475, 2016.
- [12] S. W. Wallace y S.-E. Fleten, «Stochastic Programming Models in Energy,» de *Handbooks in Operations Research and Management Science*, Elsevier, 2003, pp. 637-677.
- [13] S. Kamalinia, M. Shahidehpour y A. Khodaei, «Security -constrained expansion planning of fast-response units for wind integration,» *Electric Power System Research*, vol. 81, n° 1, pp. 107-116, 2011.
- [14] P. Xiong y C. Singh, «Optimal Planning of Storage in Power System Integrated With Wind Power Generation,» *IEEE Transactions on Sustainable Energy*, vol. 7, n° 1, pp. 232-240, 2016.
- [15] H. Heitsch y W. Römisch, «Scenario Tree Generation for Multi-stage Stochastic Programs,» de *Stochastic Optimization Methods in Finance and Energy*, New York, Springer, 2011, pp. 313-341.
- [16] Dupacová, Gröwe-Kuska y Römisch, «Scenario reduction in stochastic programming,» *Mathematical Programming*, vol. 95, n° 3, pp. 491-511, 2003.

- [17] N. Gröwe-Kuska, H. Heitsch y W. Römis, «Scenario Reduction and Scenario Tree Construction for Power Management Problems,» de *IEEE Bologna Power Tech Conference*, Bologna, 2003.
- [18] G. Xiong, X. Liu, D. Chen, J. Zhang y T. Hashiyama, «PSO Algorithm-Based Scenario Reduction for Stochastic Unit Commitment Problem,» *IEEJ Transaction on Electrical and Electronic Engineering*, vol. 12, n° 2, pp. 206-213, 2016.
- [19] Y. Feng y S. M. Ryan, «Scenario Reduction for Stochastic Unit Commitment with Wind Penetration,» de *PES General Meeting / Conference & Exposition, 2014 IEEE*, National Harbor, MD, USA, 2014.
- [20] J. Dupacová, «Postoptimality for multistage stochastic linear programs,» *Annals of Operations Research*, n° 56, pp. 65-78, 1995.
- [21] N. Rujerapaiboon, K. Schindler, D. Kuhn and W. Wiesemann, "Arxiv.org," 15 January 2017. [Online]. Available: <https://arxiv.org/abs/1701.04072>. [Accessed 02 October 2017].
- [22] J. C. Principe, D. Xu, Q. Zhao y J. W. Fisher III, «Information-Theoretic Learning,» *Journal of VLSI Signal Processing*, vol. 26, pp. 61-77, 2000.
- [23] R. Moreno, A. Street, J. M. Arroyo y P. Mancarella, «Planning lo-carbon electricity systems under uncertainty considering operational flexibility and smart grid technologies,» *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 375, n° 2100, p. 2016305, 2017.
- [24] H. Heitsch y R. Werner, «Scenario Reduction Algorithms in Stochastic Programming,» *Computational Optimizations and Applications*, vol. 24, n° 2/3, pp. 187-206, 2003.
- [25] M. Albadi y E. El-Saadany, «Overview of wind power intermittency impacts on power systems,» *Electric Power Systems Research*, vol. 80, n° 6, pp. 627-632, 2010.
- [26] J. Dupačová y V. Kozmík, «SDDP for multistage stochastic programs: preprocessing via scenario reduction,» *Computational Management Science*, vol. 14, n° 1, pp. 67-80, 2016.
- [27] A. Inzunza, R. Moreno, A. Bernales y H. Rudnick, «CVaR constrained planning of renewable generation with consideration of system inertial response, reserve services and demand participation,» *Energy Economics*, vol. 59, pp. 104-117, 2016.
- [28] S. BW, Density estimation for statistics and data analysis, London: Chapman and Hall, 1986.
- [29] J. Dupacová y V. Kozmík, «Structure of risk-averse multistage stochastic programs,» *OR Spectrum*, vol. 37, n° 3, pp. 559-582, 2014.
- [30] J. Vega, Generación de series sintéticas de recursos renovables variables para estudios de operación y planificación de sistemas eléctricos., Santiago: Universidad de Chile, 2016.
- [31] Comisión Nacional de Energía, «Informe de proyecciones de precios de combustible 2016-2031,» Santiago, 2016.
- [32] Centro de Cambio Global UC, «Estimaciones de costo y potencial de abatimiento de emisiones de gases de efecto invernadero para diferentes escenarios futuros,» Santiago, 2011.
- [33] CDEC SIC, «Estudio de previsión de Demanda 2015-2035,» Santiago, 2015.
- [34] Coordinador Eléctrico Nacional, «Proyecciones de largo plazo,» [En línea]. Available: http://cdec2.cdec-sing.cl/pls/portal/cdec.pck_web_coord_elec.sp_pagina?p_id=5047#. [Último acceso: 14 12 2017].

- [35] Proyecto Estrategia de Expansión de las Energías Renovables; Ministerio de Energía, Chile; Deutsche Gesellschaft für, «Energías Renovables en Chile. El potencial eólico, solar e hidroeléctrico de Arica a Chiloé,» Santiago, 2014.
- [36] Universidad Técnica Federico Santa María, «Estudio de contribución de las ERNC al SIC al 2025. Potencial de Biomasa en Chile,» Santiago, 2008.
- [37] Ministerio de Energía; Departamento de geofísica, Universidad de Chile, «Explorador Eólico,» [En línea]. Available: <http://walker.dgf.uchile.cl/Explorador/Eolico2/>. [Último acceso: 15 12 2017].