



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL
DEPARTAMENTO DE INGENIERIA MATEMATICA

PROGRESSIVE HEDGING APLICADO A COORDINACION HIDROTERMICA

TESIS PARA OPTAR AL GRADO DE MAGISTER EN GESTION DE OPERACIONES

MEMORIA PARA OPTAR AL TITULO DE INGENIERO CIVIL MATEMATICO

ANDRES GUILLERMO IROUME AWE

PROFESOR GUIA

ANDRES WEINTRAUB POHORILLE

PROFESOR CO-GUIA:
RODRIGO PALMA BHENKE

MIEMBROS DE LA COMISION:
ALEJANDRO JOFRE CACERES
FERNANDO ORDOÑEZ PIZARRO

SANTIAGO DE CHILE

MARZO 2012

PROGRESSIVE HEDGING APLICADO A COORDINACIÓN HIDROTÉRMICA

El problema de Coordinación Hidrotérmica busca encontrar la operación óptima para un Sistema Eléctrico Mixto, combinando en la solución los efectos de las etapas futuras así como los efectos que la hidrología tiene en la operación del sistema.

Los Sistemas Eléctricos mixtos corresponden a aquellos sistemas en los que operan tanto centrales de generación hidráulicas, geotérmicas, térmicas y eólicas entre otras. Un ejemplo de estos sistemas es el Sistema Interconectado Central (SIC) chileno, que en particular tiene una alta presencia de centrales hidráulicas y térmicas, en el cual en un año promedio un cuarenta por ciento de la energía viene de fuentes hidráulicas y un sesenta de fuentes térmicas.

Desde el punto de vista de los costos de operación, estas dos fuentes de generación presentan importantes diferencias. Mientras los costos variables de una central hidráulica son bajos, los de una central térmica son más altos debido al combustible que requiere para la generación de energía. Por otro lado las centrales hidráulicas tienen costos de inversión más elevados que las centrales térmicas.

Otra diferencia entre estas tecnologías es que las centrales hidráulicas son capaces de generar energía de acuerdo a la cantidad de agua que reciben de sus afluentes o que son capaces de almacenar (en embalses o estanques de regulación). Esta característica las hace dependientes del clima, en particular de la hidrología de una determinada zona geográfica. Debido a que no es posible predecir la hidrología, ésta se modela como una variable aleatoria.

El problema de Coordinación Hidrotérmica busca la manera óptima de operar un sistema mixto en el mediano y largo plazo. Tiene una naturaleza estocástica, debido a la incertidumbre presente al modelar la operación de las centrales hidráulicas. Corresponde a un problema de gran escala que incorpora muchos elementos; centrales de generación, redes de distribución, centros de consumo y restricciones técnicas y ambientales.

En el presente trabajo se desarrollan y aplican metodologías de programación estocástica para la resolución de un problema de Coordinación Hidrotérmica, en particular para un sistema de generación mixto. Se estudian técnicas de descomposición para problemas estocásticos que permitan trabajar con problemas de gran escala y se trabajan métodos de generación y selección de escenarios hidrológicos con el objetivo de representar de manera adecuada las componentes estocásticas del problema.

La metodología utilizada para la resolución de este problema se basa en el algoritmo Progressive Hedging (PH). En este trabajo se busca resolver un problema de planificación eléctrica a través de PH. Sobre este algoritmo se desarrollan una serie de ajustes de acuerdo a las características especiales del problema. También se realizan comparaciones con las técnicas que se utilizan actualmente para resolver este problema y se analizan las ventajas que ofrece PH para este problema en particular.

El problema de Coordinación Hidrotérmica corresponde a un problema cuadrático debido a que cuenta tanto con función objetivo cuadrática como con algunas restricciones cuadráticas. La función de costos de las centrales térmicas es modelada de manera cuadrática, así como también las pérdidas de energía en las líneas de transmisión son cuadráticas. PH es un algoritmo de descomposición por escenarios que entrega soluciones exactas para programas convexos. Funciona resolviendo sucesivas veces cada escenario por separado, penalizado por desviarse de la solución promedio. Debido a su estructura, es de naturaleza fácilmente paralelizable.

Dentro de los resultados se logra solucionar el problema para una serie de instancias, incluyendo instancias de tamaño real del Sistema Interconectado Central chileno. Además se realizan comparaciones entre PH y SDDP, otro de los métodos de solución del problema, mostrando las ventajas y desventajas que PH ofrece.

Finalmente se concluye que PH ofrece buenas posibilidades como metodología de solución para el Problema de Coordinación Hidrotérmica. Si bien actualmente no es competitivo, en el futuro, se pueden desarrollar implementaciones basadas en computación paralela que puedan ser competitivas con las técnicas actuales de resolución.

A mi familia

Índice general

1. Introducción	1
1.1. Estado del Arte	1
1.2. Stochastic Dual Dynamic Programming (SDDP)	3
1.3. Organización del contenido	3
2. Marco Teórico	4
2.1. Árboles y Escenarios	6
2.2. No-Anticipatividad	8
2.3. Formulaciones Extendidas	10
2.4. Progressive Hedging	11
2.4.1. Intuición	11
2.4.2. Pseudo-código	12
2.4.3. Caracterización	12
2.5. Contexto Eléctrico	14
2.5.1. Centrales	14
2.5.2. Sistemas Eléctricos	14
3. Modelo	15
3.0.3. Generación de Escenarios	15
3.1. Clustering de Escenarios	19

3.1.1.	<i>K</i> -means	20
3.1.2.	Error de Asignación a Clusters	21
3.2.	Modelo	22
3.2.1.	Conjuntos	22
3.2.2.	Parámetros	23
3.2.3.	Variables	24
3.2.4.	Restricciones	25
3.2.5.	Función Objetivo	27
4.	Algoritmo PH	28
4.1.	Pseudo-código extendido	28
4.2.	Ajustes y mejoras	30
4.2.1.	Acelerar la resolución de cada subproblema	30
4.2.2.	Paralelizar el algoritmo	30
4.2.3.	Linealizar el término cuadrático proximal	30
4.2.4.	Ajustar el parámetro ρ de penalización	32
4.2.5.	Criterio de término	33
4.2.6.	Bundling	33
5.	Implementación	35
5.1.	Implementación	35
5.1.1.	Instancia SIC	36
5.1.2.	Instancias Menor Tamaño	36
6.	Resultados	37
6.1.	Problema General	37
6.1.1.	Instancia SIC	37

6.1.2. Instancia SIC con Bundling	39
6.2. Comparaciones	40
7. Conclusiones y Problemas Abiertos	43
7.1. Trabajo futuro	45
8. Glosario	46
A. Descomposición de Benders	47
B. Tipos de Centrales de Generación	49
Bibliografía	51

Índice de figuras

2.1. Escenarios por separado indicando conjuntos de información (izquierda) y árbol de escenarios (derecha).	7
2.2. Relación entre Árbol de decisión y variables por escenario.	8
3.1. Afluentes Maule 1960-1969 (cms: cubic meter per second)	16
3.2. Afluentes Maule 10 años (cms: cubic meter per second)	17
3.3. Crecimiento Exponencial de Escenarios	18
3.4. Error Asignación Clusters	21
4.1. Esquema de una linealización por tramos.	31
4.2. Bundling de Escenarios	33
6.1. Convergencia PH	38
6.2. Convergencia PH + Bundling	40
6.3. Histograma F.O.	41

Índice de cuadros

3.1. Algunas Hidrologías Laguna del Maule	16
3.2. Crecimiento número de escenarios	19
4.1. Mejor ρ según instancia.	32
6.1. Convergencia PH	38
6.2. Convergencia PH con Bundling	39
6.3. Convergencia PH con Bundling	41

Capítulo 1

Introducción

En un sistema con generación hidroeléctrica se busca utilizar la energía de los caudales afluentes y la que está almacenada en los embalse para, de esta manera, disminuir los costos asociados con el uso de combustible de las unidades termoeléctricas. Pero, la disponibilidad de energía hidráulica está limitada tanto por la capacidad de almacenamiento en los embalses como por la magnitud de los caudales. Si al momento de realizar la planificación operativa del sistema no se tienen en cuenta estos elementos, el resultado será que en el futuro no se encontrarán disponibles las reservas de los embalses llevando al sistema a incurrir en costos altos, racionamiento o incluso fallas de abastecimiento. Debido a esto se genera una dependencia intertemporal con los costos que el sistema experimentará.

Esta dependencia intertemporal lleva al desarrollo de un modelo de planificación multi-etapa. Interesa conocer el costo que tendrá la energía en el futuro al momento de tomar decisiones. Por esto, se desea contar con información sobre el futuro que permita incorporar esta información a las decisiones que se adoptan para el presente.

Además, debido a que no es posible contar con pronósticos seguros sobre la disponibilidad de agua en futuro (en forma de caudales afluentes a las distintas centrales hidroeléctrica), el problema es modelado como estocástico. Lo que termina con un modelo acoplado en el tiempo y además estocástico.

1.1. Estado del Arte

El problema de Coordinación Hidrotérmica [15] corresponde a un problema de optimización estocástico representado en su equivalente determinístico como un árbol de escenarios. A través de este modelo es posible calcular la operación óptima del sistema por medio de la resolución de un problema de optimización de gran escala, equivalente a resolver la formulación extensiva del problema.

Al aumentar el número de etapas de planificación el número de escenarios crece de manera exponencial. Una manera de abordar este problema es introducir las restricciones

de no-anticipatividad (las decisiones adoptadas en una cierta etapa solo pueden depender de las realizaciones de las variables aleatorias hasta esta etapa) y luego relajarlas. De esta manera se logra descomponer el problema por escenarios.

Este modelo puede representar la realidad con distintos grados de exactitud, dependiente del número de variables con que se trabaje. En el caso del problema de forma extensa el número de variables crece de manera exponencial o en progresión geométrica con el número de etapas y de manera proporcional o en progresión aritmética con el número de escenarios para cada etapa del problema. Es por esto que la resolución del problema en su forma extensiva resulta un desafío computacional para instancias con un gran número de etapas y/o escenarios por etapa.

Desde un punto de vista metodológico existen diferentes maneras de abordar el problema. Un diseño uninodal, es decir, uno en el que se modela un sistema en el que toda la energía se genera y consume en un mismo punto (es decir, no hay una red de distribución), ofrece una aproximación más simple al problema. En cambio, un modelo multi-nodal, en el cual, si se describe en detalle la red o sistema de transmisión, pero con un significativo aumento en la complejidad del problema. En ambos casos, y principalmente debido a la componente estocástica se ha vuelto necesario utilizar técnicas de descomposición para abordar el problema. Estas técnicas, por lo general, hacen uso de la estructura escalonada de las restricciones. Debido al interés que este problema tiene para los organismos reguladores de un sistema eléctrico así como para las compañías generadoras hay una amplia variedad de modelos y técnicas de resolución que se han utilizado a lo largo de los últimos años. Algunos de ellos corresponden a:

- Modelo de Gestión Óptima del Laja (GOL), modelo de coordinación hidrotérmica uninodal y mono-embalse basado en programación dinámica estocástica. [15]
- Modelo de Operación Mensual del Sistema Interconectado Central (OMSIC), modelo de coordinación hidrotérmica uninodal y multi-embalse basado en programación dinámica estocástica. [10]
- Modelo de Programación de Largo Plazo (PLP), modelo de coordinación hidrotérmica multi-nodal y multi-embalse basado en programación dinámica dual estocástica. [15]
- Modelo Operación de Sistemas Eléctricos (OSE2000), modelo de coordinación hidrotérmica multi-nodal y multi-embalse basado en programación dinámica dual estocástica. [15]
- Modelo SEEDS desarrollado en el Centro de Energía FCFM, modelo de coordinación hidrotérmica multi-nodal y multi-embalse basado en programación dinámica dual estocástica. [15]

Se puede ver como cada uno de estos modelos utiliza programación dinámica estocástica (PDE) o bien programación dinámica dual estocástica (SDDP) [12]. Correspondiendo esta última a una aplicación estocástica de la Descomposición de Benders [2] y [1] (de ahí lo

“dual”, ver Anexo A) en su versión anidada, en la cual los subproblemas son resueltos mediante programación dinámica. Esta metodología fue desarrollada en los años 80 para el sistema eléctrico de Brasil y desde entonces se ha vuelto muy popular en países o regiones que cuentan con una importante participación de la hidroelectricidad en su matriz energética. ¹

1.2. Stochastic Dual Dynamic Programming (SDDP)

El método Stochastic Dual Dynamic Programming (o SDDP) desarrollado por Pereira y Pinto [12] corresponde a un algoritmo para resolver problemas multi-etapa de programación estocástica. Este algoritmo construye soluciones factibles en un ambiente de programación dinámica usando una sub-aproximación de una función (convexa) de costos futuros. Esta aproximación es construida a partir de cortes de Benders, que se obtienen de resolver sub-problemas asociados a escenarios específicos. Estas soluciones pueden ser revisadas luego mediante simulación para determinar su calidad y comparadas con cotas inferiores que determinan los criterios de parada para el algoritmo. Si el costo esperado de una solución factible está lo suficientemente cerca de la cota inferior, SDDP termina y entrega la solución factible. SDDP ha sido usado en problemas de planificación eléctrica y es hoy en día el método estándar para resolver estos problemas en Chile. Por este motivo, se utiliza SDDP como benchmark al trabajo realizado en esta tesis.

1.3. Organización del contenido

La tesis se desarrolla de la siguiente manera:

En el capítulo 2 se muestran los conceptos necesarios de optimización estocástica para este problema y una breve introducción a PH. Luego en el capítulo 3 se explicita el problema eléctrico en detalle junto a sus características. En el capítulo 4 se desarrolla el algoritmo, describiéndolo y mostrando los distintos ajustes y mejoras realizadas. Finalmente, en el capítulo 5, se describe y compara las dos implementaciones de la tesis. Luego, en el capítulo 6, se muestran y analizan los resultados. Para finalmente, en el capítulo 7, desarrollar las conclusiones y enunciar el trabajo futuro.

¹Centro y Sud América, Nueva Zelanda, China, California y Escandinavia entre otros otros

Capítulo 2

Marco Teórico

A continuación se presentan los conceptos relativos PH que serán utilizados a lo largo de este trabajo. Se busca formalizar las nociones de programación estocásticas más importantes y que son necesarias para el desarrollo de las implementaciones

El trabajo se enmarca en el contexto de la investigación de operaciones. En particular está relacionado con la programación estocástica aplicada a problemas de planificación con múltiples etapas. El objetivo de utilizar programación estocástica es encontrar políticas de decisiones que sean robustas frente a la incertidumbre. Es decir una planificación operativa óptima frente al desarrollo de la naturaleza, es decir, a la realización de las variables aleatorias del modelo en cada etapa. Esta planificación se manifiesta en un árbol de decisión que representa la mejor respuesta frente un futuro incierto.

El objeto de estudio es un problema de planificación eléctrica sujeto a incertidumbre externa (hidrologías). Este tipo de problema, debido a su tamaño y naturaleza estocástica, se vuelve intratable computacionalmente bajo enfoques tradicionales, luego necesita ser abordado mediante una heurística que lidie con la naturaleza del problema: variables lineales en un contexto cuadrático y múltiples etapas. Para esto se utiliza el algoritmo Progressive Hedging descrito en 2.4. Finalmente, la manera de representar la incertidumbre es a través de escenarios (ver 2.1) y la forma utilizada para medir el riesgo, es la esperanza, medida neutra al riesgo. Esta medida se utiliza porque es capaz de incorporar de manera explícita todos escenarios considerados y debido a la naturaleza práctica del problema (desde un punto de vista de tiempo de cálculo); un operador de sistema debe resolver en muchas ocasiones instancias distintas de este mismo problema.

El problema general de programación estocástica se puede formular como:

$$\begin{array}{ll} \text{mín} & \sum_{s \in S} p_s f_s(x_s) \\ \text{s.a.} & x_s \in Q_s \quad \forall s \in S \end{array} \quad (2.1)$$

En este problema existen variables de decisión x_s para cada escenario $s \in S$, la función objetivo es la esperanza (ecuación 2) y las restricciones son las propias para cada escenario (también en 2). Como se verá más adelante, en la sección relacionada a no-anticipatividad (ver 2.2), en este tipo de formulación el concepto de no-anticipatividad se encuentra de manera implícita en el problema. Esta formulación se denomina “Formulación Extendida Explícita” (ver subsección 2.3).

2.1. Árboles y Escenarios

El concepto de Árbol de Escenarios resulta muy importante a la hora de considerar problemas de programación estocástico y aporta claridad al momento de definir el concepto de no-anticipatividad (2.2 [13]). Bajo este contexto, un árbol de escenarios corresponde al conjunto de escenarios de un problema que mantiene una estructura de precedencia para cada etapa, donde un escenario corresponde a una de las posibles realizaciones de cada una de las variables aleatorias del problema. Es decir, un escenario es una tupla compuesta de un conjunto de valores numéricos determinados para todas las fuentes de incertidumbre, para todas las etapas temporales.

Ejemplo 2.1.1. Para un problema de tres etapas con dos incertidumbres, sean las variables aleatorias X_1, X_2, X_3, Y_1, Y_2 e Y_3 correspondientes a las fuentes de incertidumbres en cada etapa. Luego un escenario (s_1), se define como una realización de cada v.a. en cada etapa:

$$s_1 = (x_1, y_1, x_2, y_2, x_3, y_3) \quad (2.2)$$

Cada uno de estos escenarios tiene una probabilidad asociada. Esta probabilidad viene a su vez de las distribuciones asociadas a las variables aleatorias y en el caso más común, en el que se consideran independiente, la probabilidad del escenario corresponde al producto de las probabilidades para cada variable aleatoria.

Usualmente se cuenta con información histórica sobre alguna variable (en el caso de este trabajo se cuenta con la hidrología de los últimos años) o se estima de alguna manera la distribución de las mismas. En el caso de contar con variables aleatorias continuas, por lo habitual, se realiza una discretización para, de esta forma, obtener una estructura de árbol tradicional.

Desde un punto de vista práctico, interesan los escenarios que comparten con algún otro valores para al menos una etapa. De lo contrario, los escenarios se pueden desacoplar y no se logra una estructura tradicional de árbol.

Ejemplo 2.1.2. Para el mismo problema del ejemplo anterior, si se tienen tres escenarios s_1, s_2 y s_3 , con valores en común definidos por:

$$\begin{aligned} s_1 &= (x_1, y_1, x_2^h, y_2^h, x_3^h, y_3^h) \\ s_2 &= (x_1, y_1, x_2^h, y_2^h, x_3^m, y_3^m) \\ s_3 &= (x_1, y_1, x_2^l, y_2^l, x_3^l, y_3^l) \end{aligned} \quad (2.3)$$

Estos definen un Árbol de Escenarios como se puede apreciar en la figura 2.1. Los tres escenarios comparten los mismos valores para la primera etapa, así como el s_1 y s_2 comparten valores para la segunda etapa. Esto se puede apreciar tanto en el árbol de la derecha como en la figura con todos los escenarios de la izquierda. En este caso las elipses punteadas representan el hecho de que los escenarios encerrados son iguales para esa etapa en específico. Esto es lo que se conoce como conjunto de información.

De esta manera se puede tener una noción de cómo se incorpora la incertidumbre en el modelo. Además, se comienza a ver la relación entre la estructura de los escenarios y las

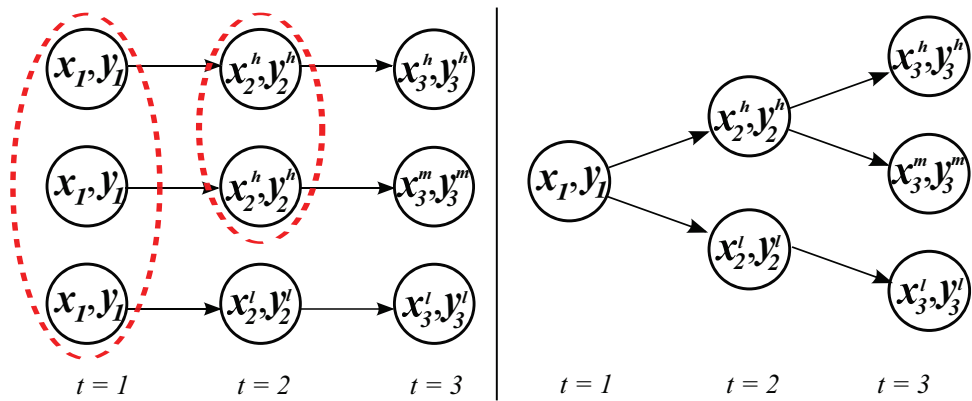


Figura 2.1: Escenarios por separado indicando conjuntos de información (izquierda) y árbol de escenarios (derecha).

variables de decisión. Es claro que si dos escenarios son iguales hasta cierta etapa (en el ejemplo s_1 y s_2 lo son hasta la segunda etapa), es porque las variables aleatorias asociadas a ellos son iguales hasta esa etapa. Esto tendrá también consecuencias sobre las decisiones que se toman para esos escenarios hasta esa etapa.

2.2. No-Anticipatividad

El principio de no-anticipatividad se hace cargo de la relación entre la estructura de un Árbol de Escenarios y las variables de decisión asociadas al problema. Fue formulado por Rockafellar y Wets (ver [13]) y corresponde a:

Definición 2.2.1. “Si dos escenarios son indistinguibles hasta cierta etapa, entonces las decisiones para esos escenarios deben ser también iguales hasta esa etapa.”

Este principio, en intuitivamente bastante natural, tiene profundas implicancias tanto teóricas como de formulación. Gráficamente, figura 2.2, la condición necesaria para que decisiones de distintos escenarios hasta cierta etapa sean las mismas, se ve como las elipses punteadas (a la derecha) que indican qué decisiones deben ser las mismas en qué etapa. Definiendo estos conjunto, “paquetes de información” que corresponde a grupos de nodos del árbol de decisión (como se puede ver en la figura 2.2 a la izquierda). Al contar con este principio se puede realizar una nueva formulación del problema, escribiendo de manera desacoplada las decisiones para los escenarios (ya sin utilizar la formulación implícita dada en un árbol) e incorporar las restricciones de no-anticipatividad.

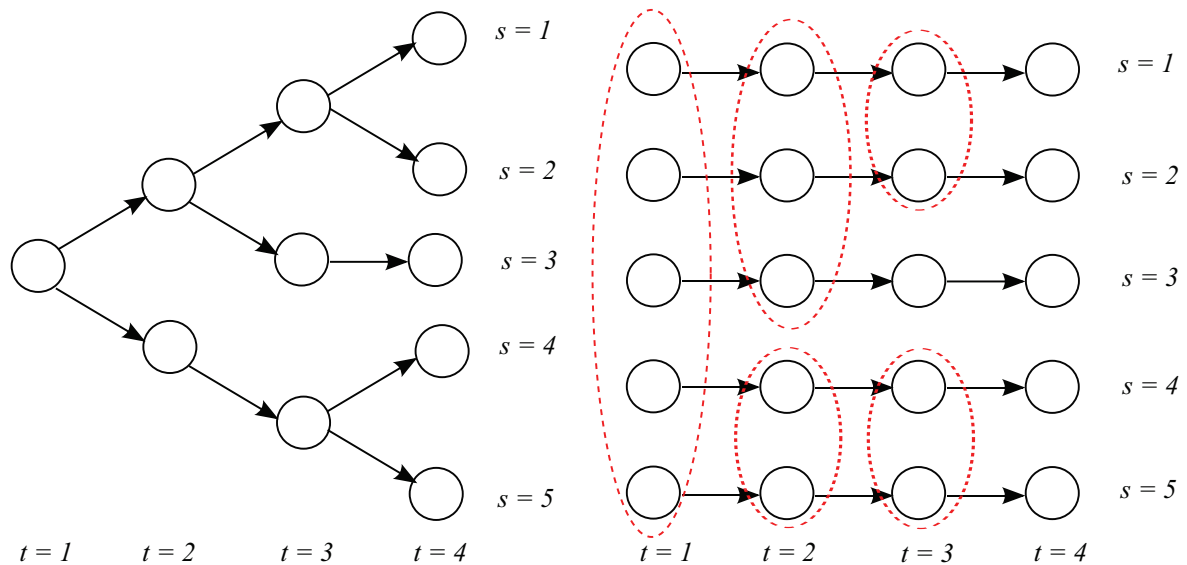


Figura 2.2: Relación entre Árbol de decisión y variables por escenario.

A continuación se introduce la notación necesaria para este tipo de formulación: El conjunto de todos los conjuntos de información es N y para facilitar la notación sus elementos (conjuntos) se indexan por las etapas del problema N_t^i . Además $N_t \cap N$ corresponde a los conjuntos de información para la etapa t . Es claro que N_t puede contener más de un conjunto de información. Algunas propiedades:

- Todo escenario pertenece a lo menos a un conjunto de información:

$$\forall s \in S \exists t \in T : s \in N_t$$

- La cardinalidad de los conjuntos de información es decreciente con respecto al tiempo:

$$|N_t^i| \geq |N_{t+1}^i| : \forall t \in T, i \quad 1 \leq |N_t| \leq |S|$$

- En la etapa inicial existe un único conjunto de información, todos los escenarios:

$$N_1 = S$$

- En la etapa final cada conjunto de información corresponde a un singleton:

$$\forall N_T^i \in N, |N_T| = 1$$

2.3. Formulaciones Extendidas

Utilizando los conceptos de no-anticipatividad descritos anteriormente, existen diferentes maneras de formular el problema de programación estocástica 2:

- Explícita: Variables por escenario con no-anticipatividad como restricción explícita
 - $\sum_{N_t \in \mathcal{N}} |N_t|$ variables por decisión, a lo más S por etapa.
 - $\sum_{t \in T} \sum_{N_t \in \mathcal{N}} (|N_t| - 1)$ restricciones por decisión para expresar la no anticipatividad.
- Utilizando la notación anterior se tiene que el problema 2 se puede reformular como:

$$\begin{aligned}
 \text{mín} \quad & \sum_{s \in S} p_s f_s(x_s) \\
 \text{s.a.} \quad & x_s \in Q_s \quad \forall s \in S \\
 & x_t^s = x_t^{s'} \quad \forall s, s' \in N_t, \forall t \in T
 \end{aligned} \tag{2.4}$$

- Compacta: Variables por conjunto de información.
 - $\sum_{N_t \in \mathcal{N}} 1$ variables por cada decisión.
 - No necesita restricciones de no-anticipatividad, las variables la implementan implícitamente.

Claramente la formulación explícita tendrá peor desempeño que la compacta al enfrentar muchos escenarios por ser de mayor tamaño, tanto en filas como columnas.

2.4. Progressive Hedging

El algoritmo Progressive Hedging PH (Rockefeller y Wets, 1989 [13]), es un algoritmo de descomposición para problemas estocásticos con escenarios y es exacto, en el sentido que converge al óptimo global para problemas convexos. Se basa en que el problema, en su formulación extendida, es intratable directamente por su gran tamaño, luego se descompone en problemas más pequeños para su resolución.

2.4.1. Intuición

La intuición detrás de esta descomposición es que si se pudiese resolver cada escenario por separado y después tomar esperanza, sólo se resolverán $|S|$ (número de escenarios) problemas simples y determinísticos. Pero intercambiar la minimización con la esperanza no es posible debido a la restricción de no anticipatividad que acopla las decisiones entre escenarios.

Ahora si se observa que para cada solución factible, debido a la restricción de no-anticipatividad, se tiene que:

$$\begin{aligned} x_s &= x_{s'} \quad \forall s, s' \in S \\ &= \mathbb{E}(x_s) \end{aligned}$$

Es decir, como todas las variables en cada nodo son iguales, éstas son iguales a su esperanza (si dos valores son iguales, en particular son iguales a su promedio). Dado esto se logra desacoplar la dependencia entre escenarios y suponiendo conocidos los multiplicadores respectivos (los penalizadores \tilde{w}_s por desviarse de la solución no-anticipatividad), se puede subir esta restricción a la función objetivo:

$$\min_{x_s \in Q_s \quad \forall s \in S} \sum_{s \in S} p_s f_s(x_s) + \sum_{s \in S} \tilde{w}_s (x_s - \mathbb{E}(x_s)) \quad (2.5)$$

Cambiando la notación y reagrupando ($\tilde{w}_s = p_s w_s$ y $\mathbb{E}(w) = 0$), se obtiene:

$$\min_{x_s \in Q_s \quad \forall s \in S} \sum_{s \in S} p_s \{f_s(x_s) + w_s x_s\} \quad (2.6)$$

De esta manera se logra que 'mín' y 'E' sean intercambiables, es decir se puede descomponer el problema en una serie de problemas más pequeños

$$\sum_{s \in S} p_s \left\{ \min_{x_s \in Q_s} f_s(x_s) + w_s x_s \right\} \quad (2.7)$$

Sólo faltará calcular, iterativamente, los multiplicadores w_s respectivos. Para asegurar la convergencia y evitar el problema de convertir los sub-problemas en no-acotados se agrega un término proximal cuadrático que fuerza a las variables a moverse en la dirección correcta:

$$\sum_{s \in S} p_s \left\{ \min_{x_s \in Q_s} f_s(x_s) + w_s x_s + \frac{\rho}{2} (x_s - \bar{x})^2 \right\} \quad (2.8)$$

Se busca que las variables sean forzadas a ir hacia el promedio \bar{x} . De lo contrario, al minimizar puede ocurrir que un multiplicador w_s vuelva no-acotada a una variable (un costo total negativo, en el caso de minimización), llevando a que el problema no converja. Ejemplo: Con objetivo lineal $f_s(x_s) = c_s x_s$ entonces $c_s^i + w_s^i < 0$ para cierta variable i minimizando.

2.4.2. Pseudo-código

Algoritmo 1: PH

```

1 Resolver cada escenario  $\min_{x_s \in Q_s} f_s(x_s)$ ;
2 Calcular la solución global en cada nodo.  $\bar{x} = \sum_{N_t: s \in N_t} p_s x_s$ ;
3 if las soluciones se parecen lo suficiente then
4   | Parar  $\|x - \bar{x}\| < \epsilon$ 
5 end
6 Actualizar la penalización.  $w = \rho(x - \bar{x}) + w_{-1}$ ;
7 Resolver cada escenario penalizado.  $\min_{x_s \in Q_s} f_s(x_s) + w_s x_s + \frac{\rho}{2} \|x_s - \bar{x}\|^2$ ;
8 Ir a 2;
```

El algoritmo busca, básicamente, resolver iterativamente cada escenario penalizando con la diferencia que tiene respecto al promedio o esperanza de todos los escenarios. De esta manera la medida de convergencia del algoritmo corresponde a la distancia euclidiana entre la solución de cada escenario y la solución global (5 en el algoritmo). Luego, el algoritmo termina cuando todas las soluciones sean suficientemente cercanas entre sí e iguales (o muy cercana) a la solución global.

2.4.3. Caracterización

Las principales características de PH (ver [13] y [6]):

- Es un algoritmo de descomposición horizontal o por escenarios. Por el contrario, descomposición vertical corresponde a descomposición por etapa. La mayoría de los enfoques anteriores para resolver el problema de planificación eléctrica corresponden a enfoques de descomposición vertical.
- No está limitado por el número de escenarios considerados, como sólo está la formulación extendida que crece mucho más rápidamente en dificultad -tamaño y tiempo de solución- al incrementar los escenarios.
- Relaja la no-anticipatividad, penalizándola en la función objetivo junto a un término proximal cuadrático.
- Posee una demostración basada en Relajación Lagrangeana Aumentada, lo que la hace exacta (converger al Óptimo global) en el caso convexo.

- Es una heurística para el caso con variables enteras (problema no convexo).
- Muy fácilmente paralelizable, basta resolver cada subproblema en distintos núcleos concurrentemente.
- Su desempeño está fuertemente ligado a lo rápido que se pueda resolver cada subproblema (cada escenario). Puesto que el resto del algoritmo básicamente es calcular unos promedios (\bar{x}) y sumas lineales sobre el mismo (w).

Sobre la aplicación y comportamiento del algoritmo en la práctica, ver el capítulo 4.

2.5. Contexto Eléctrico

2.5.1. Centrales

Con el objetivo de satisfacer la demanda por energía eléctrica de una manera eficiente y estructurada, los países han desarrollado los Sistemas Eléctricos. En estos Sistemas eléctricos se pueden identificar una serie de actores. Existen clientes o consumidores de energía y reguladores encargados de la operación del sistema, así como generadores, transmisores y distribuidores de energía eléctrica. Cada uno de estos actores cuenta con diferentes características de acuerdo a la estructura del sistema. En particular, las centrales generadores se pueden clasificar de acuerdo a la fuente energética que utilizan para generar electricidad. Para este trabajo se considera que todas las centrales del SIC corresponden a centrales hidráulicas o bien a centrales térmicas. Esto de acuerdo a sus características y estructura de costos. Para una descripción en detalle de los diferentes tipos de centrales de generación eléctrica ver [B](#).

2.5.2. Sistemas Eléctricos

El modelo tradicional de un sistema eléctrico corresponde a una estructura que cuenta con grandes centros de consumo (ciudades, grandes empresas, etc.) que son abastecidos por grandes centrales de generación. Este proceso involucra a una serie de actores (ver sección anterior sobre centrales), entre ellos se encuentran aquellos a cargo de la generación, transmisión y distribución de energía. Debido a la importancia que el sector eléctrico tiene para una economía, así como al hecho que debido a economías de escala se produzcan monopolios naturales en diferentes áreas del sistema, este cuenta tradicionalmente con un ente regulador. Es de interés para el regulador conocer la manera óptima en la que opera el sistema. Es por esto que los modelos de planificación resultan atractivos tanto para el regulador como para las compañías generadoras.

Capítulo 3

Modelo

3.0.3. Generación de Escenarios

Uno de los temas más importantes a tratar en un problema de programación estocástica es siempre la generación de buenos escenarios que sean capaces de representar de una manera apropiada la componente aleatoria del problema.

Para el problema de coordinación hidrotérmica esto se vuelve aun más importante debido a los efectos negativos que una mala selección de escenarios tiene en la calidad de las soluciones y por la importancia propia del problema.

La información histórica con la que se cuenta para la generación de escenarios corresponde a los caudales afluentes a una central. Se dispone de información histórica desde el año 1960 y la forma en la que está estructurada la información es en m^3/s promedio semanal (en un año equivalente de 48 semanas). La mayor parte de esta información se encuentra disponible en el sitio web de la Comisión Nacional de Energía (CNE) [4] y corresponden a los datos utilizados por el regulador del SIC.

El siguiente cuadro (3.1) corresponde a una selección de datos de los caudales afluentes totales al Embalse Maule. Por motivos de espacio se han dejado fuera la mayoría de los datos y se busca presentar la estructura de datos con los que se desarrolló este trabajo:

Año	Semana					
	1	2	3	...	47	48
1961	12.01	7.36	6.53	...	7.99	7.13
1962	5.96	4.60	4.18	...	11.41	10.06
1963	7.38	6.85	6.70	...	8.04	5.096517829
⋮	⋮	⋮	⋮		⋮	⋮
1999	9.68	18.22	10.33	...	3.09	3.12
2000	5.01	5.42	5.61	...	4.50	4.50

Cuadro 3.1: Algunas Hidrologías Laguna del Maule

A continuación (en la figura 3.2) se presenta la información histórica gráfica, representando 10 años (1960-1999) de información para el mismo embalse:

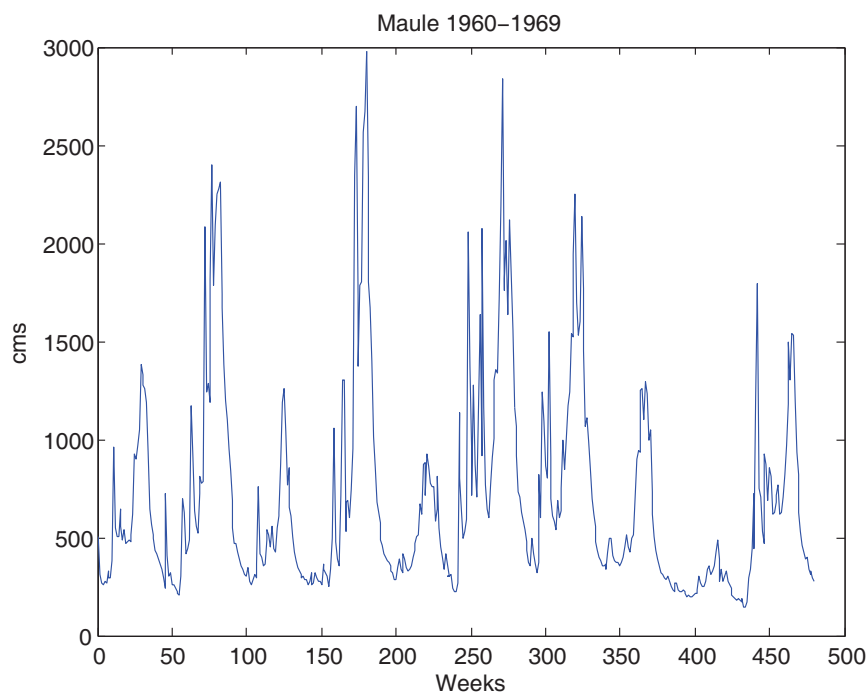


Figura 3.1: Afluentes Maule 1960-1969 (cms: cubic meter per second)

La misma información anterior, ahora separada por años, en la figura 3.2:

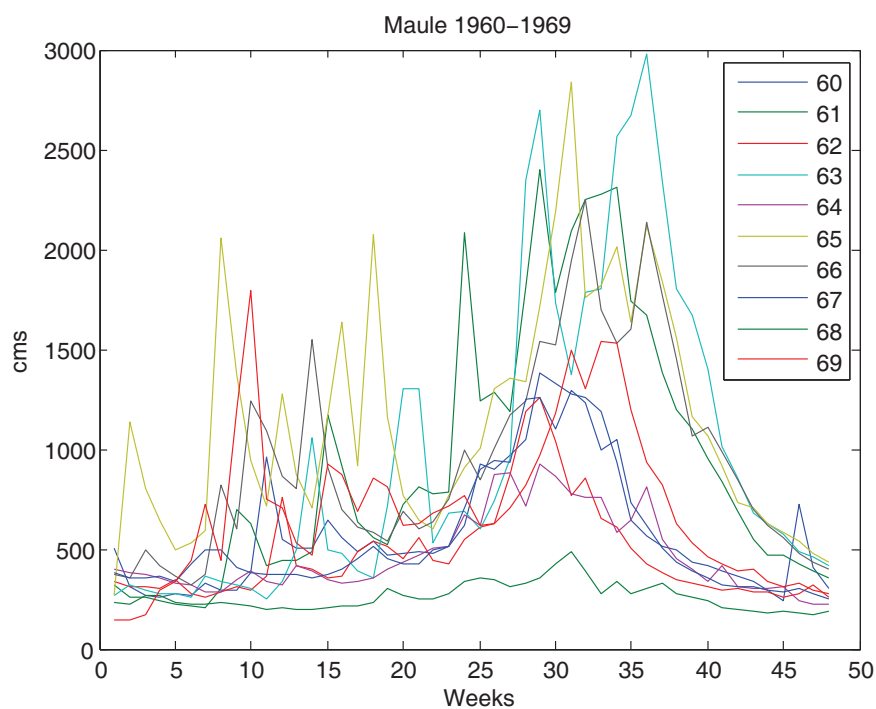


Figura 3.2: Afluentes Maule 10 años (cms: cubic meter per second)

Un escenario hidrológico puede ser perfectamente una de estas trayectorias. Es más, la manera más rudimentaria, y que se ha utilizado para enfrentar el problema de coordinación hidroeléctrica, es asumir independencia tanto temporal como espacial de los caudales afluentes. Si se desea considerar la información histórica para construir escenarios directamente a partir de ella y asumiendo que para una semana dada del futuro es equiprobable que se repitan los mismos caudales que se dieron en los 40 años anteriores se obtiene un crecimiento exponencial de variables aleatorias. La figura 3.3 pretende dar cuenta de este fenómeno.

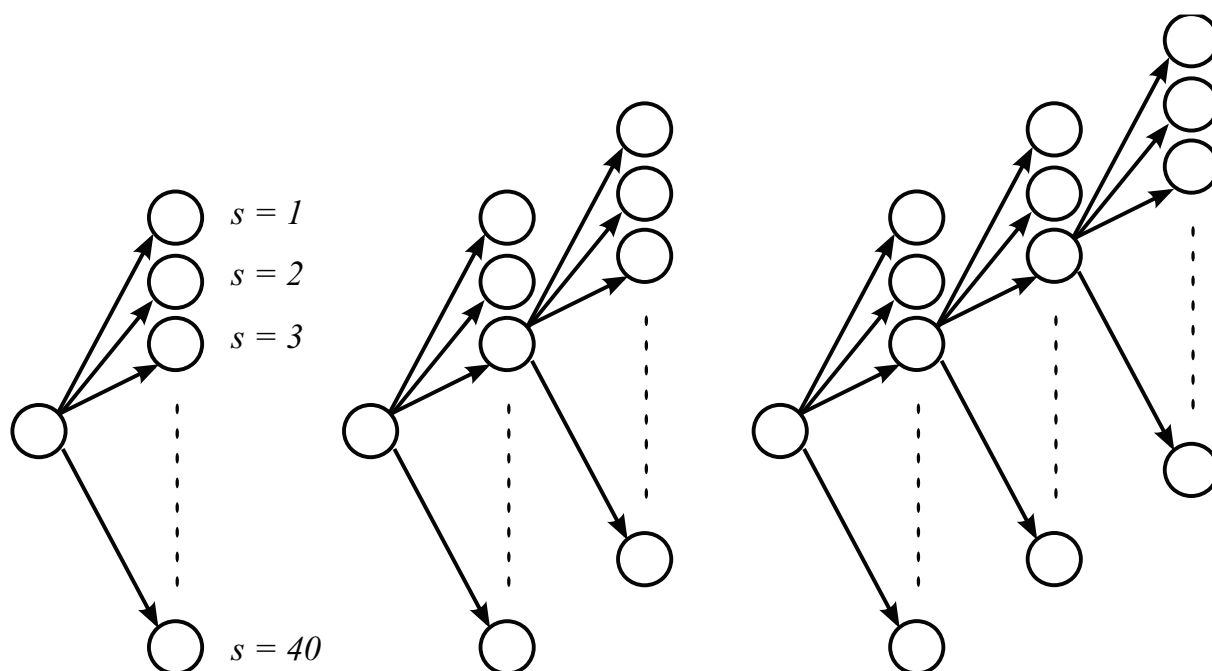


Figura 3.3: Crecimiento Exponencial de Escenarios

Lo anterior, obviamente, vuelve el problema intratable para un número relativamente pequeño de etapas. Esto, porque desde un punto de vista computacional el tener que resolver 40^t escenarios, con $t = 5, 6, 7, \dots$ es infactible debido al tamaño del árbol de escenarios.

Es por ello que se vuelve necesario contar con una manera de disminuir el número de escenarios que se utilizan para resolver el problema. En este caso el mecanismo seleccionado corresponde a la introducción de “Clustering” de escenarios, como se detalla en la siguiente parte.

3.1. Clustering de Escenarios

Frente al problema de dimensionalidad de los escenarios, una de los enfoques disponibles es encontrar un subconjunto de los escenarios originales, en el cual si sea posible resolver el problema de Coordinación Hidrotérmica. El subconjunto debe contar con una serie de requisitos para que las soluciones obtenidas tengan validez:

- Representar de una manera adecuada la estocasticidad del problema.
- Permitir encontrar soluciones en tiempo razonable.
- Mantener escenarios extremos dentro de las posibilidades.

La alternativa de Clusters ofrece una buena aproximación al primer problema. Los “Clusters” vienen a ser unas especies de “escenarios promedio” y pueden ser representativos de un gran número de escenarios que sean muy similares entre sí. Pueden, eso sí, excluir a un grupo importante de escenarios extremos, debido a que tienen una muy baja probabilidad de ocurrencia. Una sequía histórica (como la de 1968) tiene efectos considerables sobre el sistema pero una probabilidad de ocurrencia relativamente baja. Excluir este tipo de escenarios puede llevar a soluciones demasiado simples que no son capaces de capturar la naturaleza del problema real. La manera de solucionar este problema es forzar la aparición de estos escenarios como algún cluster (con una baja probabilidad) y de esta manera mantenerlos en el modelamiento.

En el cuadro 3.2 se muestra como crece el número de escenarios al considerar cuarenta posibilidades por etapa. En general, para este problema se utiliza un ramificación de cuatro escenarios por etapa. Con lo que un problema de 6 etapas tiene $4^5 = 1024$ escenarios:

	40	4	3	2
2 Etapas	1600	16	9	4
6 Etapas	10^7	1024	243	32
7 Etapas	10^9	4096	729	64
12 Etapas	10^{19}	10^7	531441	1024

Cuadro 3.2: Crecimiento número de escenarios

Es importante hacer notar que los escenarios construidos para este problema corresponden a una agregación de escenarios. En el sentido de que un escenario pasa a representar a un grupo de escenarios. Este escenario representante es uno de los escenarios original y no un promedio. De esta manera se continúa trabajando con los datos originales del problema y no con medias sobre los datos.

3.1.1. K -means

El algoritmo utilizado corresponde a una adaptación del algoritmo K -means, que busca iterativamente los K puntos medios de un set de valores. Lo siguiente corresponde al pseudo-código de este algoritmo:

Algoritmo 2: K -means Modificado

```
Inicializar: ;  
S Conjunto de escenarios;  
Seleccionar  $S_1, \dots, S_K \in S$ ,  $K$  Clusters al azar;  
1 while do  
2 | Encontrar Conjunto de Escenarios más cercano a cada cluster  $S_i$  ;  
3 | Calcular el escenario promedio  $\bar{S}_i$  dentro de este cluster  $i$ ;  
4 | Si  $S_i = \bar{S}_i$  para.;  
5 |  $S_i \leftarrow \bar{S}_i$  ;  
6 end
```

3.1.2. Error de Asignación a Clusters

Uno de los problemas enfrentados al seleccionar el enfoque de clusters es el de saber que número de clusters es necesario para representar de manera adecuada los escenarios del problema. Una manera de enfrentar este problema es definir una medida que capture el error que se comete al introducir un cluster. Para este caso se consideró:

$$E_c := \frac{|S - S_i|}{|S_i|} \quad (3.1)$$

Al realizar una estimación para este error se encuentran los valores que se muestran en la figura 3.4. De lo que se desprende que un número de 1000 escenarios representa de una manera adecuada la estructura del problema, ya que sobre este número la disminución del error es marginal:

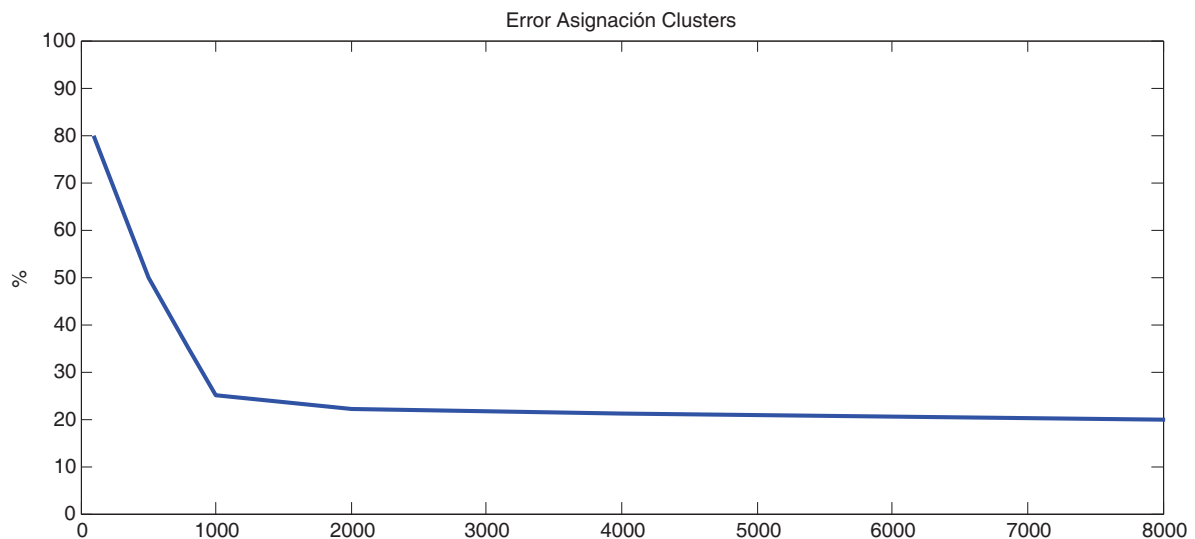


Figura 3.4: Error Asignación Clusters

3.2. Modelo

A continuación se presentan los conjuntos, parámetros, variables y función objetivo utilizados en el modelo de Coordinación Hidrotérmica. Se basa en el trabajo desarrollado por R. Torres en [15] y busca ser un modelo fácilmente comparable con otros en el área.

3.2.1. Conjuntos

Conjuntos básicos:

- T : Etapas planificación
- B : Bloques de carga
- N : Nodos de la red
- G : Generadores eléctricos
- L : Cargas o centros de consumo
- TL : Líneas de transmisión

Conjuntos sobre la red:

- Ω_i^{GN} : Conjunto de generados conectados a nodo $i \quad \forall i \in N$
- Ω_i^{LN} : Conjunto de cargas conectados a nodo $i \quad \forall i \in N$
- Ω_i^{NN} : Conjunto de nodos conectados a nodo $i \quad \forall i \in N$

Tipos de generadores:

- G_T : Generadores térmicos
- G_R : Generadores hidráulicos de embalse
- G_S : Generadores hidráulicos de pasada conectados en serie con otro generador
- G_P : Generadores hidráulicos de pasada independiente

Conjuntos sobre la red hidráulica:

- Ω_i^{RR} : Embalses aguas arriba a embals $i \quad \forall i \in G_R$
- Ω_i^{RS} : Embalses aguas arriba a central de pasada $i \quad \forall i \in G_R$
- Ω_i^{SR} : Centrales de pasada aguas arriba a embals $i \quad \forall i \in G_S$
- Ω_i^{SS} : Centrales de pasada aguas arriba a central de pasada $i \quad \forall i \in G_S$

3.2.2. Parámetros

Parámetros costos y demanda:

- α_i : coeficiente lineal función de costos generador térmico $i \quad \forall i \in G_T$
- β_i : coeficiente cuadrático función de costos generador térmico $i \quad \forall i \in G_T$
- a_i : coeficiente lineal función de costos carga $i \quad \forall i \in L$
- b_i : coeficiente cuadrático función de costos carga $i \quad \forall i \in L$
- $P_{D_i}^{t,b}$: Demanda por energía para el nodo i y bloque b en la etapa $t \quad \forall i \in N, \forall b \in B, \forall t \in T$

Parámetros líneas de transmisión:

- r_{ij} : Coeficiente línea de transmisión $(i, j) \quad \forall (i, j) \in TL$
- $\bar{f}_{(i,j)}$: Flujo máximo para línea de transmisión $(i, j) \quad \forall (i, j) \in TL$
- $\underline{f}_{(i,j)}$: Flujo mínimo para línea de transmisión $(i, j) \quad \forall (i, j) \in TL$

Parámetros conversión energía (para centrales hidráulicas):

- ρ_i^S : Factor de conversión caudal a MW para central de pasada en serie $i \quad \forall i \in G_S$
- ρ_i^P : Factor de conversión caudal a MW para central de pasada independiente $i \quad \forall i \in G_P$
- ρ_i^{R1} : Factor de conversión 1 caudal a MW para central de embalse $i \quad \forall i \in G_R$
- ρ_i^{R2} : Factor de conversión 2 caudal a MW para central de embalse $i \quad \forall i \in G_R$
- κ_i^{R1} : Factor de conversión 1 nivel de embalse a caudal filtrado para central de embalse $i \quad \forall i \in G_R$
- κ_i^{R2} : Factor de conversión 2 nivel de embalse a caudal filtrado para central de embalse $i \quad \forall i \in G_R$
- κ_i^{R1} : Factor de conversión 2 nivel de embalse a caudal filtrado para central de embalse $i \quad \forall i \in G_R$
- $\delta_i^{t,b}$: Factor de conversión nivel de embalse a caudal vertido para central de embalse i en el tiempo $t \quad \forall i \in G_R$

Parámetros restricciones de generación:

- \bar{v}_i : Nivel máximo para embalse $i \quad \forall i \in G_R$
- \bar{p}_i : Potencia máxima para generador $i \quad \forall i \in G$
- \underline{p}_i : Potencia mínima para generador $i \quad \forall i \in G$
- \bar{q}_{T_i} : Caudal máximo a turbinar para central de embalse $i \quad \forall i \in G_R$
- \bar{q}_{T_i} : Caudal máximo a turbinar para central de pasada en serie $i \quad \forall i \in G_S$
- \bar{p}_{P_i} : Caudal máximo a turbinar para central de pasada independiente $i \quad \forall i \in G_P$

Parámetros estocásticos:

- $a_i^{t,b}$: Caudal afluente a central hidráulica $i \quad \forall i \in G_R \cup G_S \cup G_P$

3.2.3. Variables

p_{itb}^T	:	Potencia generador térmico i en tiempo t y bloque b $\forall i \in G_T$
p_{itb}^R	:	Potencia central de embalse i en tiempo t y bloque b $\forall i \in G_R$
p_{itb}^S	:	Potencia central de pasada en serie i en tiempo t y bloque b $\forall i \in G_S$
p_{itb}^P	:	Potencia central de pasada independiente i en tiempo t y bloque b $\forall i \in G_P$
p_{itb}^U	:	Potencia no suministrada a nodo i en tiempo t y bloque b $\forall i \in N$
q_{itb}^{RT}	:	Caudal turbinado embalse i en tiempo t y bloque b $\forall i \in G_R$
q_{itb}^{RS}	:	Caudal vertido embalse i en tiempo t y bloque b $\forall i \in G_R$
q_{itb}^{RF}	:	Caudal filtrado embalse i en tiempo t y bloque b $\forall i \in G_R$
v_{itb}	:	Nivel embalse i en tiempo t y bloque b $\forall i \in G_R$
q_{itb}^{ST}	:	Caudal turbinado central en serie i en tiempo t y bloque b $\forall i \in G_R$
q_{itb}^{SS}	:	Caudal vertido central en serie i en tiempo t y bloque b $\forall i \in G_R$

3.2.4. Restricciones

Suministrar demanda en cada nodo de la red para cada tiempo y bloque:

$$\sum_{j \in \Omega_i^{GN}} p_{itb}^G + \sum_{j \in \Omega_i^{LN}} p_{itb}^U + \sum_{j \in \Omega_i^{NN}} f_{ij}^{tb} + \frac{r_{ij}(f_{ij}^{tb})^2}{2} \geq \sum_{j \in \Omega_i^{LN}} P_{D_j}^{tb} \quad \forall i \in N \quad (3.2)$$

Restricciones de flujo para cada línea:

$$\underline{f}_{(i,j)} \leq f_{ij}^{tb} \leq \overline{f}_{(i,j)} \quad \forall (i,j) \in TL \quad (3.3)$$

Restricciones de generación para cada central:

$$0 \leq \underline{p}_i \leq p_{itb} \leq \bar{p}_i \quad \forall i \in G \quad (3.4)$$

Restricciones para cargas:

$$0 \leq p_{itb}^U \leq P_{D_i}^{t,b} \quad \forall i \in L \quad (3.5)$$

Restricciones de caudales para embalses:

$$p_{itb}^R - (\rho_i^{R1} - \rho_i^{R2} v_{itb}) q_{itb}^{RT} = 0 \quad \forall i \in G_R \quad (3.6)$$

Restricciones conversión caudales embalses:

$$q_{itb}^{RF} - \kappa_i^{R1} v_{itb} = \kappa_i^{R2} + \kappa_i^{R1} v_{it(b-1)} \quad \forall i \in G_R \quad (3.7)$$

Restricciones conservación de flujo embalses:

$$\begin{aligned} q_{itb}^{RT} + q_{itb}^{RS} + q_{itb}^{RF} - \sum_{k \in \Omega_i^{RR}} (q_{ktb}^{RT} + q_{ktb}^{RS} + q_{ktb}^{RF}) - \sum_{k \in \Omega_i^{SR}} (q_{ktb}^{ST} + q_{ktb}^{SS}) + \frac{v_{itb}}{\delta_i^{t,b}} \\ = \frac{v_{it(b-1)}}{\delta_i^{t,(b-1)}} + a_{Ri}^{tb} \quad \forall i \in G_R \end{aligned} \quad (3.8)$$

Restricciones límite embalses:

$$0 \leq v_{itb} \leq \bar{v}_i^{t,b} \quad \forall i \in G_R \quad (3.9)$$

Restricciones caudal máximo y mínimo turbinado en embalses:

$$0 \leq q_{itb}^{RT} \leq \bar{q}_{T_i} \quad \forall i \in G_R \quad (3.10)$$

Restricciones caudal vertido en embalses:

$$0 \leq q_{itb}^{RS} \quad \forall i \in G_R \quad (3.11)$$

Restricciones caudal filtrado en embalses:

$$0 \leq q_{itb}^{RF} \quad \forall i \in G_R \quad (3.12)$$

Restricciones conversión caudales centrales de pasada en serie:

$$p_{itb}^S - \rho_i^S q_{itb}^{ST} = 0 \quad \forall i \in G_S \quad (3.13)$$

Restricciones conservación de flujo centrales de pasada en serie:

$$\begin{aligned} q_{itb}^{ST} + q_{itb}^{SS} - \sum_{k \in \Omega_i^{RS}} (q_{ktb}^{RT} + q_{ktb}^{RS} + q_{ktb}^{RF}) - \sum_{k \in \Omega_i^{SS}} (q_{ktb}^{ST} + q_{ktb}^{SS}) \\ = a_{S_i}^{tb} \quad \forall i \in G_S \end{aligned} \quad (3.14)$$

Restricciones caudal turbinado centrales de pasada en serie:

$$0 \leq q_{itb}^{ST} \leq \bar{q}_{T_i} \quad \forall i \in G_S \quad (3.15)$$

Restricciones caudal vertido centrales de pasada en serie:

$$0 \leq q_{itb}^{SS} \quad \forall i \in G_S \quad (3.16)$$

Restricciones conversión caudal centrales de pasada independiente:

$$p_{itb}^P = \min \{ \bar{p}_i^P, \rho_i^P a_{P_i}^{tb} \} \quad \forall i \in G_P \quad (3.17)$$

3.2.5. Función Objetivo

$$\sum_{t=1}^T \sum_{b=1}^B \left(\sum_{i=1}^{G_T} (\alpha_i p_{itb}^T + \beta_i (p_{itb}^T)^2) + \sum_{i=1}^L (a_i p_{itb}^U + b_i (p_{itb}^U)^2) \right) \quad (3.18)$$

Capítulo 4

Algoritmo PH

En el presente capítulo se desarrolla el algoritmo PH en detalle para el problema de Coordinación Hidrotérmica así como la manera en la que este se comparte y los ajustes y modificaciones implementados para mejorar la resolución del problema eléctrico. La información contenida en esta parte es complementaria a lo desarrollado en [2.4](#).

En general la idea de los ajustes y modificaciones es acelerar el proceso de convergencia, adecuación de los parámetros del algoritmo así como linealización de ciertos términos permite una mayor velocidad de conversión. El trade-off que se presenta es que una solución rápida puede no ser siempre óptima ya que al acelerar la convergencia se puede forzar el que el algoritmo lleve la solución a un punto sub-óptimo.

En este caso, y debido a que las restricciones no-convexas han sido omitidas, se enfrenta un problema cuadrático convexo y se puede esperar converger al óptimo (con los parámetros y tolerancia adecuados). El objetivo es encontrar soluciones adecuadas y de manera rápida.

4.1. Pseudo-código extendido

A continuación, se presenta una versión en detalle del algoritmo de PH para el caso general, incluyendo los diferentes parámetros y tolerancia necesaria. Esta versión está descrita para una formulación extendida como la de [2.4](#).

Algoritmo 3: PH

Inicializar: ;
 ε tolerancia;
 $w^0 : E(w^0) = 0$;
 $\rho^0 = 0$;
1 $k := 0, g^* := \infty$;
2 $\forall s \in S \quad x_s^k := \operatorname{argmin}_{x_s} f_s(x_s) : x_s \in Q_s$;
3 $k := k + 1$;
4 $\forall t \in T, \forall N_t \in N \quad \bar{x}_{n,t}^k := \frac{1}{|N_t|} \sum_{s \in N_t} x_{t,s}^k$;
5 $g^k := \sum_{s \in S, t \in T} \|x_{t,s}^k - \bar{x}_{n(s,t),t}^k\|$;
6 Si $g^k < g^* \vee \sum_{s \in S} f_s(x_s^k) > \sum_{s \in S} f_s(x_s^*)$, guardar mejor solución, $x^* := x^k$;
7 Si $g^k < \varepsilon \vee k > k^{max}$, ir a 13.;
8 Si $k \leq 1, \forall x_s^i \quad \rho_s^i := \rho^i(x, k, s)$;
9 $\forall s \in S, t \in T \quad w_{s,t}^k := \rho(x_{t,s}^k - \bar{x}_{n(s,t),t}^k) + w_{s,t}^{k-1}, (w^{(0)} = \vec{0})$;
10 $\forall s \in S \quad x_s^k := \operatorname{argmin}_{x_s} f_s(x_s) + \sum_{s \in S, t \in T} [w_{s,t}^k x_{t,s} + \frac{\rho}{2} \|x_{t,s} - \bar{x}_{n(s,t),t}^k\|^2] : x_s \in Q_s$;
11 Ir a 3;
12 Usando x^* como hotstart con las mismas variables fijadas, resolver la formulación extendida compacta: $\min_x \sum_{s \in S} f_s(x_s) : (x \in Q)$;

Se debe notar que:

- En (4), se utilizan elementos con la misma probabilidad luego los pesos son iguales.
- El paso (5) corresponde al cálculo de la convergencia del algoritmo, es decir, de la distancia euclidiana entre cada escenario y su óptimo estimado del nodo correspondiente. También se introduce una función $n : S \times T \rightarrow N$ para ubicar cada escenario con su nodo correspondiente en cada etapa. Esto es una alternativa simple a usar una estructura de árbol con nodos padres e hijos. Así mismo se utiliza por primera vez el subíndice n para los promedios que indica un paquete de información determinado.
- En (7) se adiciona una nueva familia de variables superindexadas con * indicando que son la mejor solución obtenida hasta el momento por el algoritmo. Esto se debe a la naturaleza heurística que no siempre converge o mejora el valor de la función objetivo, con variables enteras, sino que va saltando de solución en solución.
- En (9) se presenta el penalizador ρ como función dependiente de todas las variables, la iteración y el escenario. Esta varía según el tipo de ρ utilizado y se trata con detención en 4.2.4.
- Finalmente los pasos (1) y (3) son, respectivamente, la inicialización y actualización del iterador.

4.2. Ajustes y mejoras

Los temas investigados son:

- Acelerar la resolución de cada subproblema.
- Paralelizar el algoritmo.
- Linealización del término cuadrático proximal.
- Ajustar el parámetro ρ de penalización.
- Criterio de término.
- Bundling

4.2.1. Acelerar la resolución de cada subproblema

Esto se logra a través de tres formas; fortaleciendo el modelo determinístico (o subproblema), usando la solución de la iteración anterior como punto de partida para cada subproblema (hotstart) y ajustando la tolerancia en las diferentes iteraciones del algoritmo.

El uso de hotstarts es una estrategia ampliamente utilizada y ya fue documentada en ???. El caso del problema eléctrico no es distinto.

4.2.2. Paralelizar el algoritmo

PH cuenta con una naturaleza que lo hace fácilmente paralelizable. Cada escenario, al ser independiente de los otros, se puede resolver en un núcleo y luego reunir las soluciones de cada uno de los escenarios para así poder recalcular los nuevos promedios y penalizaciones.

Coopr cuenta con la posibilidad de realizar implementaciones paralelas de manera fácil, pero estas tienen la dificultad que, para un número bajo de procesadores o núcleos, los tiempos de cómputo de los promedios y penalizaciones exceden el ahorro de resolver los escenarios en paralelo.

4.2.3. Linealizar el término cuadrático proximal

La única desventaja de PH al resolver cada subproblema es que, aún cuando éstos son pequeños, es necesario agregarles el término proximal cuadrático de la penalización

(como se ve en 4.1) que complica bastante los subproblemas (paso 11 en 12). Esto puede aumentar en un orden de magnitud el tiempo de solución.

$$\sum_{s \in S, t \in T} [w_{s,t}^k x_{t,s} + \frac{\rho}{2} \|x_{t,s} - \bar{x}_{n(s,t),t}^k\|^2] \quad (4.1)$$

Este término es necesario para la convergencia del algoritmo, logrando que cada variable de cada escenario se acerque al promedio del nodo respectivo. También evita que los subproblemas se vuelvan no acotados debido al término w_x , que representa la no-anticipatividad y puede tomar signo positivo o negativo (dependiendo de la relación de la variable con el promedio), donde uno de estos permitiría a la variable respectiva crecer indefinidamente, dependiendo de si se minimiza o maximiza, si es que se trata de una variable no acotada. Pero este término tiene un sentido de asegurar convergencia y si es reemplazado por otro que asegure lo mismo no se pierde su sentido. Luego, si este término se puede reemplazar por otro que asegure convergencia al mismo tiempo que disminuya el tiempo de solución, esto tendrá un valor al momento de correr el algoritmo.

Dada esta necesidad de rapidez, si se mira el término con detención, se aprecia que tan sólo son variables multiplicadas por si mismas (ej.: xx o yy) pero no hay términos cruzados (mismo ej.: xy). Luego una aproximación lineal por tramos (piece-wise linealization) como la mostrada en la figura 4.1 puede ser fácilmente utilizada.

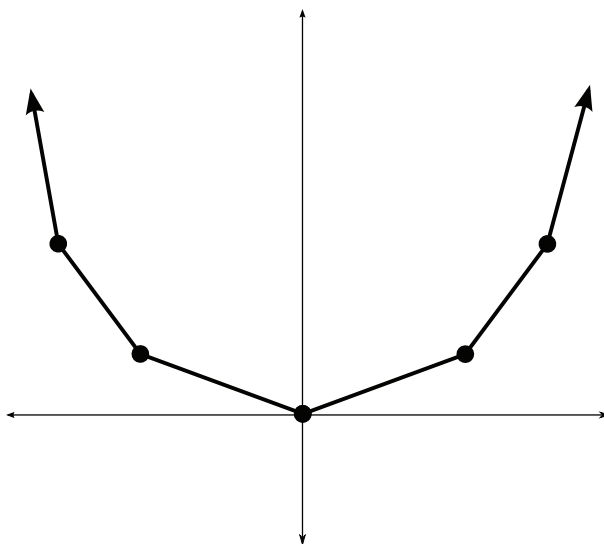


Figura 4.1: Esquema de una linealización por tramos.

Esta linealización fue implementada en este trabajo (de la manera como se muestra en la ecuación 4.2), considerando que es importante actualizar las pendientes y los puntos de quiebre de la linealización en cada escenario en cada iteración debido a que el punto mínimo de la curva representada es el promedio \bar{x} y este puede cambiar en todas las

iteraciones. También es importante considerar que es necesario que las variables linealizadas tengan cotas propias debido a que las pendientes calculadas acotan el problema sólo finitamente y no como la curva original $\rho(x - \bar{x})^2$ que siempre puede ser mayor que wx cuando $x \rightarrow \pm\infty$, entonces puede darse que en un problema de cierta variable no acotada al linealizar se vuelve no acotado.

$$x^2 - 2x\bar{x} + \bar{x}^2 \longrightarrow 2x - 2x\bar{x} + \bar{x}^2 \quad (4.2)$$

4.2.4. Ajustar el parámetro ρ de penalización

El parámetro ρ (ver (2.4.2) paso 4 o (4.1) paso 9) es el ponderador relativo entre forzar la no anticipatividad y cumplir el término proximal como se puede apreciar en la función objetivo del subproblema modificado (ver 14 paso 5 o 12 paso 11). Este parámetro es muy sensible dentro del algoritmo como ya fue documentado por Mulvey y Vladimirov a finales de los 80's. También debido a como se construye el ponderador w , ρ puede interpretarse como el precio fijado por desviarse de la solución ponderada, y así w tiene la natural interpretación de ser el cobro extra dentro de la función objetivo para que al subproblema le convenga tomar la solución no anticipativa.

Luego la premisa de mejora al modificar este parámetro es que si se puede encontrar un mejor sistema de precios que mejore la relación de las variables entonces el algoritmo convergerá más rápido y mejor. El punto de partida de las estrategias testeadas nuevamente es Watson y Woodruff en [6].

Las estrategias estudiadas en este trabajo fueron, ρ : *FIX*, *CP*, *SEP*, *Taylor* y combinaciones de ellas. Las primeras tres aparecen en [6] y la última *-Taylor-* fue desarrollada por F. Badilla del DII. Finalmente, la selección de ρ estuvo subordinada al evitar la convergencia anticipada de las variables debido a la linealización extrema de sólo 2 tramos y a la rapidez de su cálculo. Estos requisitos se cumplieron al tomar un ρ fijo y muy pequeño como se muestra en el cuadro siguiente 4.1

N° de Escenarios	ρ
18	10^{-1}
64	10^{-1}
128	10^{-2}
256	10^{-2}
512	10^{-2}
1024	10^{-3}

Cuadro 4.1: Mejor ρ segun instancia.

Como se aprecia en la tabla anterior, el tamaño de ρ óptimo crece con el número de escenarios. Esto ocurre debido a que con mayor cantidad de escenarios, hay más discrepancias entre los escenarios, haciendo la convergencia más difícil y disminuyendo la cantidad de variables fijadas por iteración. Luego para no necesitar más iteraciones (y más tiempo) se aumenta el tamaño de ρ .

4.2.5. Criterio de término

El criterio de término del algoritmo (paso 8 en 12), si bien incluye terminar luego de alcanzar la tolerancia mínima para aceptar un problema como convergido (es decir cuando todos los escenarios se parecen suficientemente a su óptimo correspondiente del árbol), en la práctica esto no se da puesto que toma mucho tiempo y el valor de la función objetivo empeora al forzar la convergencia. Luego el criterio que se usó fue realizar un número razonable de iteraciones de tal manera de converger rápidamente. El criterio obtenido empíricamente fue $\epsilon = 0,01$ y $k^{max} = 30$.

4.2.6. Bundling

Un Bundle de escenarios corresponde a un conjunto de escenarios que comparte a lo menos un conjunto de información del árbol de escenarios. Si se quiere, en lenguaje de Teoría de Grafos, corresponde a un sub-árbol de escenarios. En la figura 4.2 se muestra gráficamente un grupo de tres Bundles para un árbol de cinco escenarios:

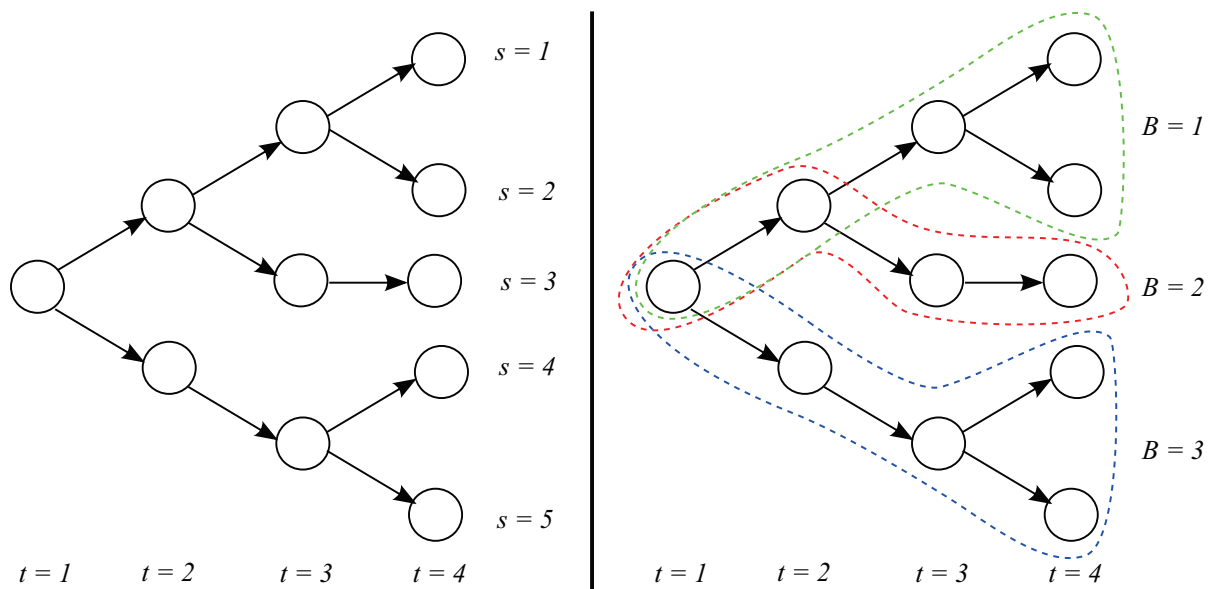


Figura 4.2: Bundling de Escenarios

Los Bundles corresponden a macro-escenarios y se pueden resolver a través de su equivalente determinístico. En este caso, si PH resuelve un Bundle como su equivalente

determinístico, el Bundle pasa a ser un nuevo escenario pero con una cantidad mayor de variables y restricciones que en el caso anterior.

La idea detrás de este enfoque es que si los escenarios se resuelven más rápido que el tiempo que se toma PH en realizar el resto de sus pasos (cálculo de promedios, actualización de penalizaciones, etc.) entonces conviene reducir el número de veces que se resuelve cada escenario. Es decir, en el caso sin Bundle, se resuelven muchas iteraciones de problemas fáciles y en el caso con Bundles se resuelven menos iteraciones en las que los macro-escenarios son más complejos y por lo mismo requieren un tiempo mayor para ser resueltos.

Capítulo 5

Implementación

5.1. Implementación

Para realizar la implementación de PH se trabajó con el Profesor David Woodruff de University of California, Davis. El Profesor Woodruff junto a un grupo de investigadores de Sandia National Labs han desarrollado software específicos para la resolución de problemas de optimización estocástico. Estos software han sido programados en el lenguaje Python y se encuentran dentro de la colección de paquetes denominada Coopr [5].

El lenguaje para modelar estos problemas se denomina Pyomo y el paquete para problemas estocásticos corresponde a PySP. Este paquete permite una adecuada interacción con los problemas estocásticos y permite separar el modelo determinístico de los escenarios. Con esto se pueden realizar experimentos de manera rápida y a través de un constructor de escenarios, construir instancias con el número y estructura de escenarios deseada.

El solver utilizado, tanto para correr PH como para realizar las comparaciones con SDDP, corresponde a Cplex 12.1. Para la construcción de escenarios se utilizó Matlab y para la construcción de escenarios para PySP se programaron una serie de scripts en Python 2.7.

Los datos utilizados en estos experimentos se encuentran disponibles en el sitio web de la CNE [4] y formarán parte del set de Bases de Cálculo y Archivos de Salida Informe Definitivo que esta institución elabora dos veces al año.

Las corridas para realizar comparaciones se realizaron utilizando el paquete SEEDS Study Manager del Software DeepEdit. Este corresponde a una herramienta para solucionar problemas de coordinación hidrotérmica de mediano y largo plazo y fue desarrollado en el Centro de Energía de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. También se contó con la colaboración del Centro de Energía para aprender a utilizar SEEDS y poder correr los problemas de prueba.

En los experimentos realizados se utilizaron dos tipos de instancias. En el capítulo de resultados (ver 6) se indica a cual de estos tipos corresponde cada resultado. Primero se buscó resolver una instancia completa del Sistema Interconectado Central. Luego se utilizaron instancias de menor tamaño tanto en número de variables como en escenarios. Este tipo de instancias permiten realizar un mayor número de experimentos y, en general, rescatan los análisis que se buscaban realizar.

Los escenarios utilizados en ambas instancias fueron generados utilizando los métodos descritos en la sección sobre generación de escenarios (ver 3.0.3). En los resultados del capítulo 6 se señala la cantidad de escenarios utilizados en los diferentes experimentos que se muestran. En aquellos casos en los que se realizaron comparaciones con SDDP se utilizaron los mismos escenarios para los dos métodos de solución.

Las maquinas utilizadas para correr los experimentos corresponden a:

- Estación de Trabajo, 2.2GHz AMD Athlon architecture running Linux, with 64GB of RAM.
- PC, 2.5GHz Intel architecture running Linux, with 3GB of RAM.
- 2.5GHz Intel architecture running Windows 7, with 3GB of RAM.

5.1.1. Instancia SIC

- 200 Nodos
- 161 Centrales Térmicas
- 11 Centrales de Embalse
- 10 Centrales de Pasada en Serie
- 67 Centrales de Pasada Independientes
- 293 Líneas de transmisión

5.1.2. Instancias Menor Tamaño

- 2 Centrales de Embalse
- 3 Centrales Térmicas
- 2 Nodos
- 12 Líneas de transmisión

Capítulo 6

Resultados

A continuación se presentan los resultados obtenidos en este trabajo de acuerdo a lo señalado en la parte de implementaciones 5. Se buscó encontrar dos tipos de resultados. Primero resolviendo el problema de coordinación hidrotérmica (ver 3.2) para un sistema completo (en este caso el Sistema Interconectado Central) con la mayor cantidad de escenarios posible, para, de esta manera, asegurar la correcta representación de la naturaleza estocástica del problema. En segundo lugar se buscan resultados que permitieran realizar una comparación con los métodos de solución que actualmente se utilizan (principalmente SDDP 1.1) para resolver este problema.

6.1. Problema General

Debido a la gran escala del problema para una instancia del SIC, estas corridas se realizaron en una estación de trabajo (ver detalles en 5) de Sandia National Labs. En general, las instancias del SIC no se pueden correr en un PC, excepto para aquellas con un número muy reducido de escenarios.

6.1.1. Instancia SIC

En el cuadro 6.1 se presentan los resultados de la mejor corrida obtenida para una instancia del SIC con 1024 escenarios. La primera columna indica la iteración de PH, mientras que la segunda indica la métrica de convergencia (ver 4). La tercera columna indica el valor de la función objetivo para esa iteración y, finalmente, la cuarta entrega la máxima distancia para una variable entre escenarios. Esta corrida tomó un tiempo de 2134,547 segundos y el parámetro de PH ρ se fijó como $\rho = 0,001$.

Iteración	Covergencia	F.O.	Dist
0	7972.1970	1030217.0154	2953.831
1	5735.7930	1055416.7489	2886.725
2	3627.1870	1046105.1922	2505.467
3	1908.8613	1046029.5800	245.1660
4	835.2984	1054214.3225	241.8485
5	238.7123	1053595.8283	228.1147
6	213.0512	1040838.9649	191.1531
7	126.2100	1044280.9769	17.48372
8	75.8642	1056783.4582	16.6364
9	38.2928	1044839.0181	15.0256
10	35.7202	1047737.3198	14.3834
11	26.2689	1051500.0002	13.05797
12	20.1784	1053891.1937	9.22138
13	12.5702	1050392.0321	4.4666
14	9.9072	1046146.1823	0.61

Cuadro 6.1: Convergencia PH

La información más importante en la tabla 6.1 corresponde a la velocidad de convergencia del algoritmo. La figura 6.1 muestra como evoluciona la convergencia en cada iteración de PH.

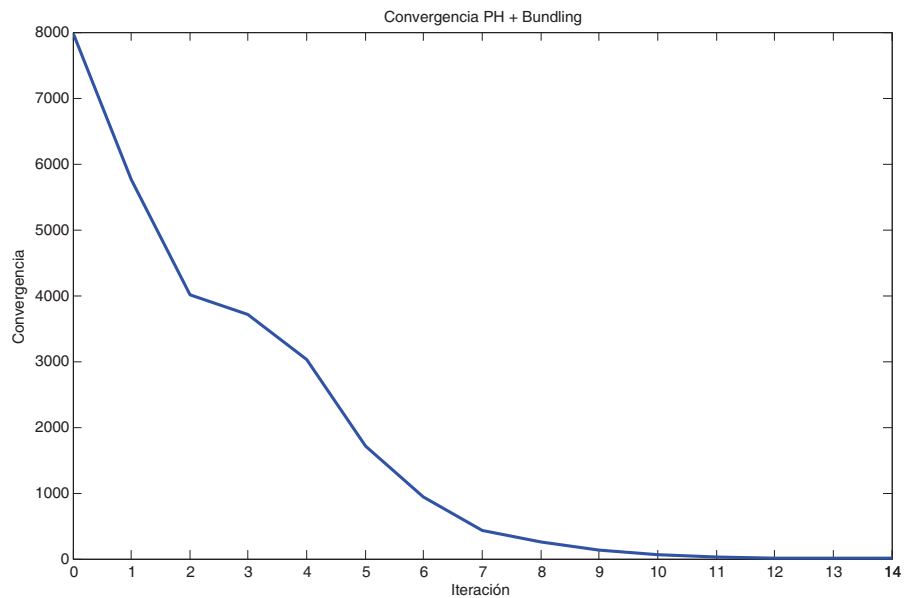


Figura 6.1: Convergencia PH

6.1.2. Instancia SIC con Bundling

En el cuadro 6.2 se presentan los resultados de la mejor corrida obtenida para una instancia del SIC con Bundling con 1024 escenarios, agrupados en 64 Bundles de 16 escenarios cada uno. Como se describe en 4, la técnica de Bundling permite convertir un grupo de escenarios en un macro-escenario y resolver el equivalente determinístico de este en cada iteración de PH, en lugar de resolver cada escenario por separado. Este enfoque es efectivo en aquellos casos en los que el tiempo que demora resolver un escenario es considerablemente menor al tiempo que demora cada iteración de PH (el cálculo de promedios, actualización de penalizaciones, etc.). La estructura de la tabla es la misma que en la tabla 6.1 en la sección anterior. Esta corrida tomó un tiempo de 1065,535 segundos y el parámetro de PH ρ se fijó como $\rho = 0,1$.

Iteración	Covergencia	F.O.	Dist
0	7972.1970	1030217.0154	4126.712
1	5765.6331	1042150.7191	2145.777
2	4004.6747	1046486.5844	1897.392
3	3711.6422	1048302.5436	1469.772
4	3022.1210	1043291.7113	1384.612
5	1706.4969	1041142.3034	107.5159
6	943.7610	1040447.5990	102.6401
7	436.3042	1044780.0011	93.3982
8	250.0928	1044955.8615	84.3696
9	134.5427	1043140.2851	77.2620
10	61.5394	1043072.3117	54.4005
11	22.4450	1047151.9215	2.5913
12	13.4511	1049160.0256	7.0648
13	7.0579	1052599.2105	1.26
14	6.4378	1052137.1244	0.571

Cuadro 6.2: Convergencia PH con Bundling

Nuevamente, la información más relevante en la tabla 6.2 corresponde a la velocidad de convergencia del algoritmo. La figura 6.2 muestra como evoluciona la convergencia en cada iteración de PH. Es importante notar que el enfoque de Bundling no solo permite resolver el escenario en un menor tiempo total, sino que además la velocidad de convergencia es mayor a la del caso anterior.

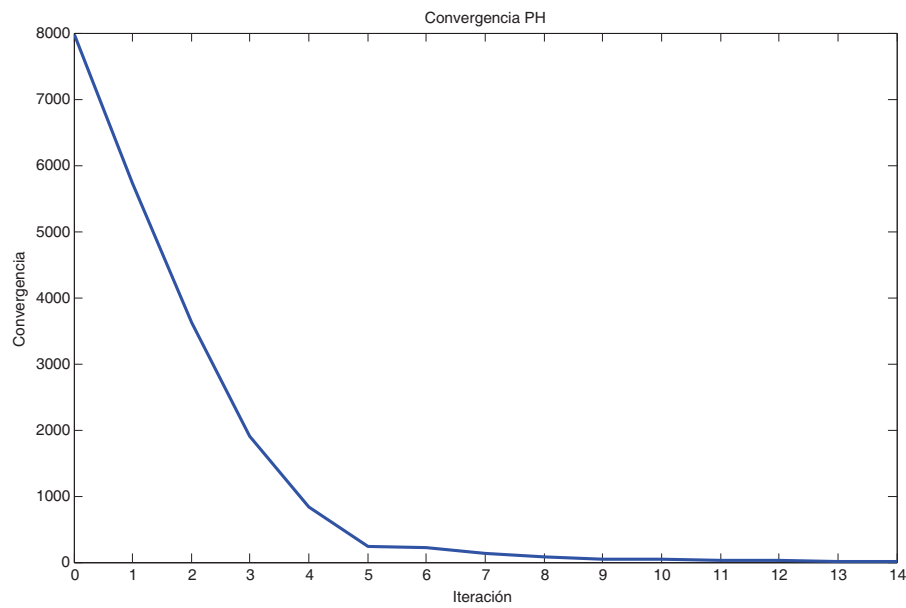


Figura 6.2: Convergencia PH + Bundling

6.2. Comparaciones

Debido a que para realizar las comparaciones se deseaba poder ambos métodos de resolución (SDDP y PH) en una misma máquina y como solo es posible resolver las instancias del SIC con PH en una estación de trabajo, se optó por correr instancias más pequeñas del problema en un PC. Además esta aproximación permite realizar un número mayor de experimentos en el mismo tiempo debido a que las instancias más pequeñas pueden ser corridas en tiempos considerablemente menores.

El número de escenarios para estas corridas corresponde a 64 y la instancia es la descrita en 5. La tabla 6.3 muestra los resultados para una corrida de PH y se compara con los resultados para el mismo problema resuelto con SDDP. Los tiempos de ejecución de SDDP son, por lo general, menores. Al mismo tiempo PH ofrece resultados más robustos en aquellos casos en los que SDDP excluye escenarios extremos en su *sampling*.

Si bien es cierto que SDDP puede solucionar este problema al realizar un gran número de *samplings* (un enfoque tipo Sample Average Approximation de [8]), PH al considerar todas los escenarios no necesita hacerlo.

La tabla 6.3 muestra la convergencia de PH para una instancia pequeña. El formato es el mismo que el usado en la sección anterior

Iteración	Covergencia	F.O.	Dist
0	1706.4969	97484.7977	701.1559
1	943.7610	121753.5802	201.4601
2	436.3042	125754.9825	92.9382
3	250.0928	134912.3377	48.6396
4	134.5427	138873.3744	26.7720
5	61.5394	143670.8365	4.4505
6	45.9	146949.5400	2.5913
7	12.10	146985.2517	4.0678
8	3.46	147834.3810	2.16
9	0.536	148004.004	0.517

Cuadro 6.3: Convergencia PH con Bundling

El tiempo de solución para esta corrida fue de 20.96 segundos.

Mientras tanto, al resolver esta corrida, SDDP (Corriendo SEEDS, ver 5) tomó un tiempo de 2,81 segundos y obtuvo como resultado un valor para la función objetivo de 147070.0107 ($< 1\%$ de diferencia).

Esto se explica porque SDDP realiza un *sampling* de escenarios y puede dejar afuera casos extremos. Al evaluar las soluciones (la decisión operativa) que se obtiene de ambos métodos en un escenario extremo seco (con valores para los caudales afluentes muy bajos), se observa como la solución de PH: 151097.6797 es menor a la de SDDP: 154125.1785. SDDP obtiene una función que es un 1,4% más alta, es decir al realizar el *sampling* subestima este escenario extremo.

En la figura 6.3 se muestra el histograma para los valores de la función objetivo al evaluar ambas metodologías en cada uno de los 64 escenarios de la instancia:

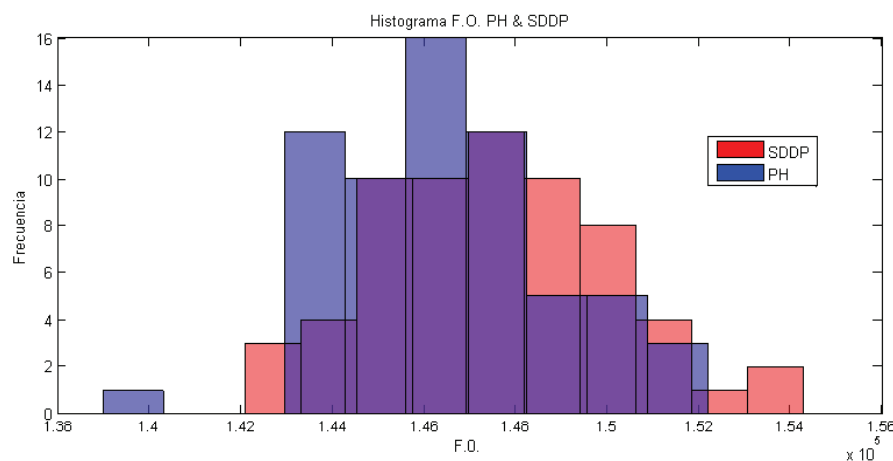


Figura 6.3: Histograma F.O.

Es importante mencionar que este histograma se obtuvo para un *sampling* en particular de SDDP. Este *sampling*, subestima los costos en el caso de escenarios con hidrologías

bajas (aquellos con costos más altos). Por lo mismo el histograma de SDDP se encuentra desplazado hacia la derecha respecto del histograma de PH. Al realizar otras corridas de SDDP se pueden obtener resultados diferentes, por ejemplo con el histograma de SDDP a la izquierda del de PH. Lo que se busca mostrar, es como la selección aleatoria de escenarios que realiza SDDP tiene efecto en su función objetivo.

Capítulo 7

Conclusiones y Problemas Abiertos

A continuación se presentan las conclusiones de este trabajo. Estas se separan en tres grupos. Primero las conclusiones relacionadas con el algoritmo de PH, luego aquellas correspondientes con la implementación del problema y la corrida de las diferentes instancias y finalmente las conclusiones relativas a la comparación entre PH y SDDP.

- Se comprobó la utilidad de PH para resolver problemas de múltiples etapas y con un alto número de escenarios.
- Se documentó y describió una serie de mejoras y ajustes al algoritmo para mejorar su desempeño en este caso particular.
- El algoritmo PH presenta una alta sensibilidad a los parámetros que se utilizan para las corridas. Es por ello que un estudio para analizar cuáles son los mejores valores de los parámetros de penalización es muy útil.
- La técnica de Bundling de escenarios ofrece una importante mejora en los tiempos de solución para este problema. El agrupar escenarios permite resolver equivalentes determinísticos más grandes y esto ayuda a disminuir el número de iteraciones que PH requiere para converger.
- Se logró resolver instancias reales del SIC para este problema de hasta 1024. Si bien esto se realizó en una estación de trabajo, se muestra que PH es capaz de resolver problemas de esta escala.
- La técnica de generación de escenarios implementada (mediante Clusters) permitió reducir la dimensionalidad del problema hasta tamaños manejables. Gracias a esta técnica fue posible mantener la estructura estocástica del problema y no perder robustez.
- Si bien PH ofrece algunas ventajas teóricas sobre SDDP, en el sentido de que resuelve problemas sin recurrir a un sampling de escenarios. Es necesario hacer notar que bajo las condiciones actuales PH no se comporta de manera competitiva. Por un lado, sin una implementación paralela, no es posible aprovechar la estructura de

PH. Pero además, y dado que los problemas de gran escala, PH no es capaz de resolverlos de una manera competitiva.

- El algoritmo PH, a diferencia de SDDP, resuelve todos los escenarios con los que se trabaja y por eso se comporta de manera más robusta frente a cualquier realización de uno de los escenarios. SDDP al realizar un sampling de escenarios puede excluir alguno y no capturar eventos extremos. Si bien este problema se corrige realizando más de un sampling, es interesante notar esta diferencia entre ambos métodos.
- La incorporación de eventos extremos corresponde a uno de los aportes más significativos que tiene el considerar programación estocástica para resolver un problema. Incluir esto eventos por medio de escenarios, aun cuando tenga una baja probabilidad de ocurrencia, ayuda a construir modelos mejor preparados para enfrentar cualquier situación.

La contribución de este trabajo corresponde principalmente a proponer una metodología alternativa a SDDP para problemas de planificación eléctrica en el Sistema Interconectado Central. Se adaptó e implementó el algoritmo PH para resolver una serie de instancias de este problema y se realizó un análisis comparativo con SDDP. Además se propone una nueva manera de abordar el problema de la generación de escenarios. Se presentaron las opciones que tiene PH para ser un algoritmo competitivo para resolver estos problemas.

7.1. Trabajo futuro

Los principales trabajos a futuro están relacionados con implementar mejores métodos de construcción de escenarios así como también de desarrollar implementaciones paralelas de PH.

- Si bien el enfoque de generación de escenarios mediante clusters fue útil para este problema, resulta claro que puede ser mejorado. En el enfoque solo se consideró como utilizar la información histórica de buena manera. Además de esto, es posible construir escenarios incluyendo información geográfica y meteorológica relevante a la hidrología de un sector. Utilizando este tipo de inputs es posible desarrollar procesos estocásticos que entreguen buenos escenarios para ser utilizados en el problema.
- La implementación paralela ofrece enormes oportunidades de mejora para PH. Si bien esta es un área más cercana a la programación que a investigación de operaciones, los resultados que se obtendrían justifican el esfuerzo. Por otro lado, técnicas de computación paralela desarrolladas en particular para PH como la el cálculo de manera asincrónico de las penalizaciones ayudan a acelerar las iteraciones del algoritmo. Es decir, si es que se resuelve cada sub-problema por separado, se abre la posibilidad de no tener que esperar que todos los escenarios estén resueltos antes actualizar las penalizaciones. La ventaja de esto es que se puede usar la información relevante de cada escenario antes e incluso resolver varias veces un escenario fácil antes que uno difícil.
- El problema de Coordinación Hidrotérmica es solo uno de los muchos problemas que existen en planificación eléctrica. Es posible aplicar PH a otra serie de problemas. En particular a aquellos problemas con decisiones de inversión que incluyan variables binarias. La experiencia que se ha obtenido al resolver problemas enteros con PH en otras áreas permite esperar buenos resultados en esta tipo de problemas también.

Desarrollar un sistema adecuado para generación de escenarios que sea capaz de representar de manera adecuada la naturaleza del sistema y al mismo tiempo permita resolver instancias de tamaño interesante, aparece como de primera necesidad. En el caso de la implementación paralela, es posible construirla en base al trabajo realizado para Coop. El principal trabajo a realizar, en este aspecto, es disminuir el overhead que se produce al momento de combinar las soluciones a los distintos escenarios.

Capítulo 8

Glosario

- CDEC: Centro Económico de Despacho de Carga, operador del sistema eléctrico chileno.
- CNE: Comisión Nacional de Energía.
- CH o PCH: Coordinación Hidrotérmica o Problema de Coordinación Hidrotérmica.
- Coopr: Paquete de software para optimización desarrollado en Sandia National Labs con aplicaciones de programación estocástica.
- PYOMO: Lenguaje de modelamiento de Coopr.
- PYSO: Paquete de programación estocástica de Coopr.
- SDDP: Stochastic Dual Dynamic Programming, método de programación estocástica para resolver problemas multi-etapa.
- SIC: Sistema Interconectado Central, sistema eléctrico de la zona centro-sur de Chile.

Anexo A

Descomposición de Benders

Si se considera el siguiente problema de optimización:

$$\begin{aligned} \min_{x,y} \quad & cx + fy \\ \text{s.a.} \quad & Ax + By \geq b \\ & y \in Y \\ & x \geq 0 \end{aligned} \tag{A.1}$$

Donde las variables x e y tienen una naturaleza diferente (en el contexto de este trabajo; x puede corresponder a las decisiones para la primera de un problema de planificación e y a las decisiones para las etapas siguientes).

En el caso de contar con valores fijos para y , se puede escribir el problema como:

$$\begin{aligned} \min_{x,y} \quad & cx \\ \text{s.a.} \quad & Ax \geq b - By \\ & x \geq 0 \end{aligned} \tag{A.2}$$

Replantando el problema inicial:

$$\min_{y \in Y} \left[fy \min_{x \geq 0} \{ |Ax \geq b - By\} \right] \tag{A.3}$$

De esta manera el dual del problema interior esta dado por:

$$\begin{aligned} \max_u \quad & (b - B\bar{y})^t u \\ \text{s.a.} \quad & A^t u \leq c \\ & x \geq 0 \end{aligned} \tag{A.4}$$

Luego, la Descomposición de Benders resuelve dos problemas de manera iterativa. Un problema maestro restringido de la forma:

$$\begin{aligned}
 & \underset{y}{\text{mín}} && z \\
 & \text{s.a.} && z \geq fy + (b - By)^t \bar{u}_i \quad i = 1 \dots I \\
 & && (b - By)^t \bar{u}_j \leq 0 \quad j = 1 \dots J \\
 & && y \in Y
 \end{aligned} \tag{A.5}$$

Además de los sub-problemas:

$$\begin{aligned}
 & \underset{u}{\text{máx}} && f\bar{y} + (b - B\bar{y})^t u \\
 & \text{s.a.} && A^t u \leq c \\
 & && x \geq 0
 \end{aligned} \tag{A.6}$$

De esta manera el Algoritmo para la Descomposición de Bender corresponde a:

Algoritmo 4: Descomposición de Bender

```

Inicializar: ;
y solución factible;
LB := -∞ ;
UB := ∞ ;
1 while UB - LB > Tolerancia do
2   Resolver sub-problema (dual) ;
3   máxu(f $\bar{y}$  + (b - B $\bar{y}$ )tu | Atu ≤ c, x ≥ 0);
4   if No acotado then
5     |  $\bar{u}$  rayo no acotado;
6     | Agregar corte factibilidad (b - B $\bar{y}$ )t $\bar{u}$  al problema maestro;
7   else
8     |  $\bar{u}$  punto extremo;
9     | Agregar corte optimalidad z ≥ fy + (b - By)t $\bar{u}$  al problema maestro;
10    | UB := mín(UB, f $\bar{y}$  + (b - B $\bar{y}$ )t $\bar{u}$ );
11  end
12  Resolver maestro: míny(z | cortes, y ∈ Y);
13  LB :=  $\bar{z}$ ;
14 end

```

Anexo B

Tipos de Centrales de Generación

Las centrales de generación eléctrica se clasifican (ver [3]) como:

- **Hidroeléctricas:** Corresponde a aquellas centrales que son movidas por agua. Pueden transformar la energía cinética de un cauce de agua (utilizando su caudal) y/o la energía potencial presente en un desnivel.
 - **Centrales de Pasada:** Aquellas que no tienen capacidad de almacenar agua y deben generar de acuerdo a la demanda y disponibilidad que tienen en un momento dado.
 - **Centrales de Embalse:** Las que si pueden almacenar agua con el objeto de generar electricidad en el futuro. De acuerdo a su capacidad se puede tener una regulación de va desde semanas hasta un año.
 - **Centrales Mixtas:** Se puede almacenar agua en pequeñas cantidades (normalmente durante el día) para poder generar en las horas de mayor demanda.
 - **Centrales de Bombeo:** Presentan un desnivel y un estanque en el nivel inferior. Durante las horas de baja demanda se bombea el agua hasta el nivel superior para ser usada durante las horas de demanda alta.
 - **Centrales Mareomotrices y Marinas:** Aquellas que utilizan los flujos de agua de mar o la presión de las olas para generar electricidad.
- **Centrales a Vapor (Térmicas):** Cuentan con una turbina que genera electricidad al ser movida por vapor de agua a alta temperatura y presión. El vapor es generado en calderas que pueden utilizar diferentes combustibles como carbón, petróleo, gas natural u otros. Debido a que necesitan calentar el vapor, estas centrales cuentan con tiempos de partida y apagado largos.
- **Centrales Geotérmicas:** Utilizan la energía térmica presente en el magma terrestre para mover una turbina.
- **Centrales Nucleares:** Su operación es similar a las centrales térmicas pero el combustible corresponde a material radioactivo y funcionan realizan fusión nuclear dentro de los reactores.

- **Turbinas a gas:** Cuentan con una turbina que es activada directamente por los gases de combustión en lugar de ser movida por vapor de agua. Presentan tiempos de encendido menores que las centrales térmicas tradicionales.
- **Centrales de ciclo combinado:** Utilizan una caldera que aprovecha los gases de escape de turbinas a gas. Esta caldera funciona luego como una central a vapor.
- **Motores Diesel:** Se utilizan en sistemas pequeños. Necesitan además del combustible, aire y agua para su refrigeración.
- **Centrales Eólicas:** Utilizan la energía cinética del viento para mover una turbina.
- **Centrales de Cogeneración:** Corresponde a las centrales que genera tanto energía eléctrica como térmica durante su proceso. Esto permite aumentar la eficiencia y se usan tradicionalmente en procesos industriales (como plantas de celulosa).
- **Centrales Solares:** Utilizan la energía solar para generar electricidad. Puede ser a través de una conversión directa en el caso de los paneles fotovoltaicos o a través de una turbina movida por vapor calentado por espejos que concentran la luz del sol.

Bibliografía

- [1] D. Bertsimas and J. N. Tsitsiklis. Introduction to linear optimization. 1997.
- [2] J.R. Birge and F. Louveaux. Introduction to stochastic programming. 1997.
- [3] W. Brokering, R. Palma, and L. Vargas. *Ñom Lüfke (El Rayo Domado) o Los Sistemas Eléctricos de Potencia*. Ed. Prentice Hall-Pearson Educación, 2008.
- [4] Comisión Nacional de Energía CNE. <http://www.cne.cl>.
- [5] Sandia National Labs Coopr. <https://software.sandia.gov/coopr/>.
- [6] David L. Woodruff Jean-Paul Watson and David R. Strip. Progressive hedging innovations for a class of stochastic resource allocation problems.
- [7] Roger J-B Wets Jean-Paul Watson and David L. Woodruff. Scalable euristics for stochastic programming with scenario selection. *Technical Report, Discrete Math and Complex Systems Department, Sandia National Laboratories Albuquerque*, 2008.
- [8] A.J. Kleywegt, A. Shapiro, and T. Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.
- [9] A. Løkketangen and D.L. Woodruff. Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics*, 2(2):111–128, 1996.
- [10] Eduardo Minder. Modelación hidrológica en el sistema interconectado central. 2007.
- [11] E.C. Finardi M.L.L. dos Santos, E.L da Silva and R.E.C. Gonçalves. Practical aspects in solving the medium-term operation planning problem of hydrothermal power systems by using the progressive hedging method. *Electrical Power and Energy Systems*, 2009.
- [12] MVF Pereira and L. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1):359–375, 1991.
- [13] R.J-B.Wets Rockafellar, R.T. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 1991.
- [14] Ubuntu. Sistema Operativo GNU/Linux. <http://www.ubuntu.com>.

-
- [15] Rigoberto Torres. Modelo de coordinación hidrotérmica multinodal y multiembalses para sistemas eléctricos longitudinales en el mediano plazo. 2006.