

UN MODELO DE COORDINACIÓN HIDROTÉRMICA DE CORTO PLAZO MEDIANTE DESCOMPOSICIÓN ANIDADA ESTOCÁSTICA DE BENDERS

Andrés Ramos

Santiago Cerisola

Universidad Pontificia Comillas

Alberto Aguilera 23

28015 Madrid, España

andres.ramos@iit.upco.es

Resumen

Este artículo presenta un modelo de coordinación hidrotérmica de corto plazo con un alcance de varios meses formulado como un problema de programación lineal estocástica multietapa. Para su resolución se emplea el método de descomposición anidada estocástica de Benders. Se han analizado y desarrollado diversas técnicas y mejoras en la resolución de los subproblemas y en la implantación del algoritmo de descomposición anidada para reducir su tiempo de ejecución. El modelo ha sido implantado en el lenguaje de modelado GAMS y utiliza el optimizador CPLEX para la resolución de los problemas de optimización. Se presenta un caso de estudio de muy gran tamaño derivado del sistema eléctrico peninsular español.

Palabras clave

Programación lineal estocástica multietapa. Árbol de escenarios. Corte virtual. Base previa.

1 Introducción

Bajo el actual marco regulatorio la gestión de los recursos de generación es responsabilidad de las propias empresas eléctricas, porque pueden determinar los precios y cantidades que cada generador (o porfolio, en el caso de los subsistemas hidráulicos) oferta en el mercado. En particular, los recursos hidroeléctricos juegan un papel relevante en los diversos mercados por su naturaleza intrínseca de recursos sin coste variable de operación asociado y por la flexibilidad en su uso.

Un modelo de coordinación hidrotérmica de corto plazo ayuda en la toma de decisiones sobre la producción a realizar con energía hidráulica para un alcance típico de varios meses a un año, teniendo explícitamente en cuenta la estocasticidad de las aportaciones. El modelo plantea la gestión hidrotérmica como un problema de *optimización lineal estocástica* de muy gran tamaño (como el caso de estudio que se presenta), que no puede ser resuelto directamente con los optimizadores actuales (al menos bajo el entorno de lenguajes de modelado). Su resolución se hace mediante la técnica de *descomposición anidada estocástica de Benders* (también denominada programación dinámica dual estocástica), ver [3], [4] y [5], que permite la resolución de un problema de gran tamaño con una estructura en escalera en la matriz de restricciones mediante la solución iterativa de otros problemas de menor tamaño. Este método evita el problema de dimensionalidad asociado a la programación dinámica al centrarse en la evaluación de los estados en las sucesivas iteraciones alrededor de los estados de interés.

El modelo se ha implantado en lenguaje de modelado algebraico GAMS 2.50 [2] con el optimizador CPLEX 6.0 para la resolución de los subproblemas. Este lenguaje permite la implantación rápida, eficiente y compacta de problemas de optimización.

2 Caracterización del sistema

Existe una decisión preliminar sobre qué sistema representar que conviene analizar brevemente:

- ***Representación del sistema propio***
Se entiende por sistema propio el de una compañía eléctrica cualesquiera. Tiene la dificultad del modelado de la demanda, que corresponde a la producción que la compañía prevé colocar en el mercado (sobre el cual se tiene de momento poca experiencia).
- ***Representación del sistema peninsular español***
Es más conservadora en el modelado ya que continúa la experiencia de la operación centralizada. Tiene la desventaja de que se necesitan hacer suposiciones sobre costes o comportamiento de la competencia.

En el modelo se ha tomado la segunda opción, representar la operación del sistema eléctrico peninsular español. Se trata de un sistema hidrotérmico de gran tamaño con numerosas centrales hidráulicas.

2.1 División temporal y demanda

El *alcance* temporal del modelo es de varios meses a un año, dividido en *periodos*, *subperiodos* y *niveles* de carga de duración variable. Una división adecuada es la telescópica, con mayor detalle temporal al comienzo y periodos más largos al final del alcance. La demanda es cronológica con valores constantes para cada nivel de carga.

2.2 Subsistema térmico

Los grupos térmicos se representan individualmente con su potencia de operación dividida en dos bloques, mínimo técnico y resto. La indisponibilidad fortuita se representa de manera determinista reduciendo la potencia de operación proporcionalmente a la tasa de indisponibilidad. La curva de consumo de combustible se toma como una recta. Se contemplan rampas de subida y bajada de los grupos entre niveles de demanda consecutivos.

2.3 Subsistema hidráulico

El modelado de los subsistemas hidráulicos es especialmente difícil por dos razones. Una es la propia complejidad de las redes hidráulicas de las cuencas y la otra es la indisponibilidad de datos relativos a las características y aforos de las centrales de la competencia. Se supone que en el futuro únicamente se dispondrá de datos precisos de las centrales y embalses de la propia compañía. Los de los competidores sólo serán conocidos de manera aproximada y requerirán un modelado diferente.

En las centrales propias se consideran explícitamente las redes hidráulicas que enlazan embalses, la dependencia de la potencia producida del caudal y del nivel de reserva del embalse. En la versión actual del modelo, las curvas de rendimiento de las centrales se linealizan alrededor del punto de trabajo dando lugar a un problema de optimización lineal. En el futuro, se abordará la utilización de aproximaciones poligonales con variables binarias para la selección del segmento adecuado. El algoritmo de descomposición utilizado permite la aplicación natural a este tipo de problemas binarios mixtos. Sin embargo, una implantación ingenua haría que el tiempo de ejecución resultara un orden de magnitud superior o, visto de otra manera, que el tamaño del problema de optimización fuera sensiblemente menor para un mismo tiempo.

Las centrales de la competencia se agregan por cuencas y se representan como centrales equivalentes con embalses de gestión independiente.

Para todas las centrales las aportaciones son estocásticas y se organizan en un árbol de escenarios como el de la figura 1, compuesto por 43 escenarios agrupados en 8 periodos. Cada nodo representa un escenario dado por una aportación y una probabilidad y tiene un ancestro y uno o varios descendientes. En la figura 1, el primer periodo tiene dos aportaciones y el segundo y tercero sólo tienen una.

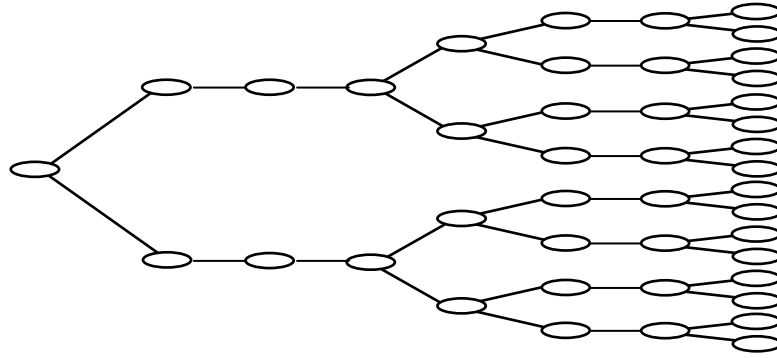


Figura 1. Árbol de escenarios de aportaciones hidráulicas estocásticas.

3 Modelo matemático

El problema de optimización lineal estocástica multietapa se define genéricamente de esta manera:

- ***Función objetivo***

Minimización del valor esperado de los costes variables de explotación para el conjunto de escenarios de aportaciones hidráulicas.

- ***Restricciones INTERPERIODO***

Aquellas en las que intervienen variables de dos periodos consecutivos. El problema es separable por periodos excepto por estas ecuaciones. Corresponden al movimiento de las reservas para cada escenario: el nivel de reserva en un embalse en un periodo más las aportaciones menos la producción o vertido de dicho embalse es igual al nivel de reserva en el siguiente periodo.

$$R_{ph}^{\omega} + A_{ph}^{\omega} - M_{ph}^{\omega} = R_{p+1h}^{\omega}$$

donde R_{ph}^{ω} es el nivel de reserva al comienzo del periodo p en el embalse h para el escenario ω , A_{ph}^{ω} son las aportaciones y M_{ph}^{ω} es el movimiento neto del nivel de reserva por la producción, vertido o bombeo de las centrales conectadas a él.

- **Restricciones INTRAPERIODO**

Aquellas en las que sólo intervienen variables de un periodo. Son las relativas a la operación de los grupos de generación (definición del movimiento neto de un embalse, limitación en la energía bombeada y turbinada dentro del periodo, límites de operación de cada grupo, seguimiento de los niveles de reserva dentro del periodo, balance entre generación y demanda, reserva rodante en el sistema, etc.)

- **Variables de decisión**

Son las variables de generación del sistema (producción de los grupos en cada nivel, turbinación y bombeo en cada embalse) así como las variables de estado asociadas (niveles de reserva de los embalses).

3.1 Descomposición anidada determinista

A continuación se presenta la formulación matemática genérica del problema de *programación lineal determinista multietapa* de P etapas que se resuelve por descomposición anidada determinista y en el siguiente apartado se introduce la estocástica.

$$\begin{aligned} \min_{x_p} \quad & \sum_{p=1}^P c_p^T x_p \\ & B_{p-1} x_{p-1} + A_p x_p = b_p \quad p = 1, \dots, P \\ & x_p \geq 0 \end{aligned}$$

donde las matrices A_p y B_p (con $B_0 \equiv 0$) y los vectores b_p y c_p son parámetros conocidos y x_p es el vector de variables.

La descomposición anidada es una extensión natural del algoritmo de descomposición de Benders bietapa [1] para problemas multietapa. Básicamente, consiste en aplicar de forma recursiva la descomposición del problema bietapa. Se considera la etapa 1 como problema maestro y las etapas 2,...,P como el subproblema. De nuevo se aplica la descomposición hasta llegar a la etapa $P-1$ como problema maestro y la etapa P como subproblema. En cada momento sólo se necesita resolver un problema maestro en las etapas 1 a $P-1$ o como subproblema en la etapa P . El algoritmo rompe las restricciones interperiodo y resuelve aisladamente la operación del sistema en cada periodo. El subproblema genérico que ha de resolverse en la etapa p es el siguiente:

$$\begin{aligned} \min_{x_p, \theta_p} \quad & c_p^T x_p + \theta_p \\ & A_p x_p = b_p - B_{p-1} \bar{x}_{p-1} \quad : \pi_p \\ & \theta_p + \bar{\pi}_{p+1}^T B_p x_p \geq f_{p+1} + \bar{\pi}_{p+1}^T B_p \bar{x}_p \quad : \eta_p \\ & x_p \geq 0 \end{aligned}$$

siendo $\theta_p \equiv 0, B_0 \equiv 0, \pi_{p+1} \equiv 0, \eta_{p+1} \equiv 0$ condiciones adicionales de la formulación, θ_p es la función de recursos de la etapa p (es decir, el coste futuro desde $p+1$ hasta P), f_{p+1} indica un coste futuro real en la etapa $p+1$, $\bar{x}_{p-1}$ es la propuesta enviada por el ancestro del problema p y π_p, η_p son las variables duales de las restricciones.

3.2 Descomposición anidada estocástica

El problema lineal estocástico multietapa de P etapas se puede formular como:

$$\begin{aligned} \min_{x_p^{\omega_p}} \quad & \sum_{p=1}^P \sum_{\omega_p \in \Omega_p} p_p^{\omega_p} c_p^{\omega_p T} x_p^{\omega_p} \\ B_{p-1}^{\omega_p} x_{p-1}^{\omega_p} + A_p^{\omega_p} x_p^{\omega_p} &= b_p^{\omega_p} \quad p=1, \dots, P \\ x_p^{\omega_p} &\geq 0 \end{aligned}$$

siendo $B_0^{\omega_1} \equiv 0$, Ω_p representa el conjunto de los posibles valores de ω_p en la etapa p cada uno con probabilidad de ocurrencia $p_p^{\omega_p}$. Estas probabilidades son, en general, dependientes de los valores tomados en la etapa anterior. El problema determinista equivalente (PDE) se obtiene al enumerar explícitamente todos los estados del sistema. En la descomposición anidada estocástica multietapa cada problema es maestro de sus problemas descendientes en el periodo siguiente y subproblema del ancestro en el periodo anterior. En cada iteración del algoritmo se realiza una pasada resolviendo cada subproblema desde el primer al último periodo y viceversa. Hacia adelante cada subproblema ancestro decide el estado final de cada embalse y lo transfiere a los subproblemas descendientes. Hacia atrás se envía la información dual necesitada para la formación de los cortes de Benders, que son aproximaciones lineales exteriores de los costes futuros en función de los niveles finales de los embalses, ver figura 2.

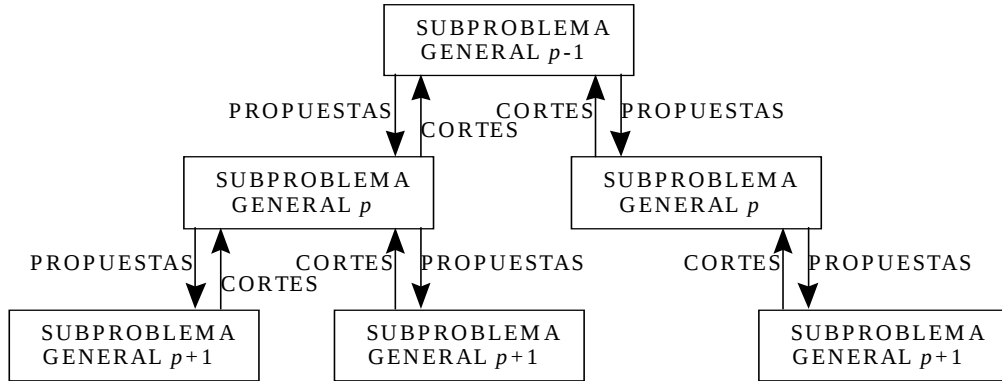


Figura 2. Esquema del algoritmo de descomposición anidada estocástica de Benders.

La formulación de un subproblema genérico para la etapa p es la siguiente:

$$\begin{aligned} \min_{x_p^{\omega_p}, \theta_p^{\omega_p}} \quad & c_p^T x_p^{\omega_p} + \theta_p^{\omega_p} \\ A_p^{\omega_p} x_p^{\omega_p} &= b_p^{\omega_p} - B_{p-1}^{\omega_p} \bar{x}_{p-1}^{A(\omega_p)} & : \pi_p^{\omega_p} \\ \sum_{k \in D(\omega_p)} p_{p+1}^k \left\{ \bar{\pi}_{p+1}^{kT} B_p^k x_p^{\omega_p} \right\} + \theta_p^{\omega_p} &\geq \sum_{k \in D(\omega_p)} p_{p+1}^k \left\{ f_{p+1}^k + B_p^k \bar{x}_p^{\omega_p} \bar{\pi}_{p+1}^{kT} \right\} & : \eta_p^{\omega_p} \\ x_p^{\omega_p} &\geq 0 \end{aligned}$$

siendo $\theta_p^{\omega_p} \equiv 0, B_0 \equiv 0, \pi_{p+1}^{\omega_p} \equiv 0, \eta_{p+1}^{\omega_p} \equiv 0$.

Se formula el método *monocorte*, donde todos los descendientes de un subproblema forman un único corte agregado para devolver dicha sensibilidad.

4 Avances en la implantación del algoritmo

Se han implantado mejoras en dos aspectos importantes, unidos entre sí y necesarios para la resolución de problemas de gran tamaño, que garantizan estabilidad y robustez en el comportamiento numérico y en la convergencia del algoritmo: resolución eficiente de los subproblemas y mejoras en el algoritmo de descomposición

4.1 Resolución eficiente de los subproblemas

Para la resolución eficiente y numéricamente estable de los subproblemas se han seguido estas líneas de actuación:

a) *Reformulación del subproblema*

El criterio general es disminuir el número de restricciones y/o elementos de la matriz de restricciones. Se prueban diferentes formulaciones matemáticas del mismo problema de optimización hasta conseguir aquella con menor tiempo de resolución.

b) *Escalación natural de variables y restricciones*

Para garantizar el buen comportamiento numérico tanto las variables como las restricciones se escalan a unidades físicas naturales alrededor de 1.

c) *Selección automática del método de optimización más adecuado*

Para la resolución de problemas lineales existen dos familias de algoritmos: punto interior y simplex, primal o dual. El método de punto interior es apropiado para problemas de tamaño medio con estructura cuasi vacía de la matriz de restricciones mientras que el método simplex resulta adecuado en la realización de análisis de sensibilidad (es decir, cuando se trata de resolver problemas similares disponiendo de una solución próxima y una base previa).

Ambas características han sido aprovechadas en este modelo. Para la primera resolución de un subproblema en la primera iteración se utiliza el método de punto interior, mientras que en sucesivas resoluciones se utiliza el método simplex primal.

d) *Utilización de bases previas*

Para resoluciones sucesivas de problemas similares mediante el método simplex es conveniente aprovechar las bases de los hermanos y primos resueltos previamente.

Como ejemplo del impacto en el tiempo de optimización del aprovechamiento de una base, un problema lineal de 8000 restricciones, 10000 variables y 30000 elementos requiere 10.3, 4.4, 4.7 y 2.6 segundos¹ en sucesivas resoluciones.

Aun aprovechando el método simplex una base previa de un problema similar anterior, aproximadamente a partir de tamaño de 10000 x 10000 el método de punto interior resulta más efectivo.

4.2 Mejoras en el algoritmo de descomposición anidada

Entre las diferentes mejoras introducidas en el algoritmo de la descomposición anidada estocástica con objeto de acelerar la convergencia destacan las siguientes:

¹ Las referencias de tiempo que aparecen en el artículo se han tomado para un ordenador personal con procesador Pentium II a 233 MHz con sistema operativo Windows NT con 64 MB de memoria principal.

a) Resolución del problema determinista del valor medio (PDVM)

Antes de resolver el problema estocástico se resuelve primeramente el problema determinista con los valores medios de las aportaciones (o cualquier otro valor). Las soluciones iniciales de las variables en todos los escenarios de cada periodo son los valores óptimos de las variables del periodo correspondiente en el caso medio o determinista resuelto en primer lugar.

b) Cotas de factibilidad de las variables interperiodo

Su función es delimitar la senda factible de las variables interperiodo (niveles de reservas de los embalses) debido a la dificultad que tiene cada subproblema de ver (el coste futuro) más allá de su periodo. Estas cotas, que no restringen el óptimo, sólo garantizan la factibilidad de cualquier subproblema, mejoran la convergencia al estrechar la región factible y evitan tener que considerar cortes de infactibilidad en los subproblemas.

e) Modificación de parámetros del optimizador

En el algoritmo de descomposición para evitar expresamente la degeneración y mejorar la convergencia se especifican parámetros de control del optimizador.

c) Secuencia de barrido del árbol de escenarios

Por su sencillez de implantación y por su efectividad, se ha elegido la pasada rápida (*fast-forward*, *fast-backward*) [4] que consiste en recorrer el árbol en cada iteración desde la primera etapa hasta la última y viceversa.

d) Descomposición en superperiodos

El tiempo de recorrido de un subproblema se compone del tiempo de formulación (creación) en GAMS y del tiempo de resolución por el optimizador. Existe un equilibrio entre estos tiempos teniendo en cuenta que el tiempo de resolución depende también del método de optimización.

El algoritmo de descomposición permite el recorrido del árbol de escenarios de cualquier manera. La que mejor se adapta a la estructura del problema y su implantación en GAMS es la resolución simultánea de un nodo y sus descendientes, mientras éstos sean descendientes únicos. Para ello se detectan estas ramas de ancestros con descendientes únicos en el árbol a lo largo de varios periodos (los denominados superperiodos). Por ejemplo, en la figura 1 los periodos 2, 3 y 4 pueden formar un superperiodo. En un superperiodo solamente las ecuaciones del primer se ven afectadas por las propuestas del subproblema ancestro y las del último periodo por los cortes de los subproblemas descendientes.

La elección del recorrido del árbol con periodos o superperiodos se realiza automáticamente en función de un número preespecificado de restricciones del subproblema para lograr un tiempo de recorrido mínimo.

e) Corte virtual

Debido a la estructura específica de los problemas formulados en el problema estocástico la solución obtenida en un problema para un periodo puede ser utilizada en otro subproblema del mismo periodo. Esta solución relajada (no óptima) permite formar cortes sin necesidad de resolver los subproblemas y, en algún caso, acelerar la convergencia del algoritmo.

5 Caso de estudio

El caso de estudio representa el sistema eléctrico peninsular español con un modelado detallado del subsistema hidráulico de una compañía. Su alcance temporal es de 8 periodos con un total de 121 escenarios. Los tamaños de los problemas para este caso se presentan en la tabla 1.

	Restricciones	Variables	Elementos
Problema determinista equivalente (PDE)	827856	1325136	3356856
Problema determinista de valor medio (PDVM)	35055	55427	142174
Subproblema de menor tamaño	1018	1632	4020
Subproblema de mayor tamaño	9149	14672	36711

El tiempo de resolución del PDVM es de 80 s y el tiempo medio para un subproblema es de 4 s. El algoritmo requiere 15 iteraciones para alcanzar la convergencia con una tolerancia relativa entre cota superior e inferior del problema de $5 \cdot 10^{-3}$ y un tiempo total de resolución de 4 horas.

Si el PDE pudiera ser formulado y resuelto requeriría una cantidad de memoria principal casi 100 veces superior a la requerida por el mayor subproblema.

6 Conclusiones

En este artículo se ha presentado un modelo de coordinación hidrotérmica de corto plazo con un alcance de varios meses formulado como un problema de programación lineal estocástica multietapa de muy gran tamaño. Se ha utilizado el método de descomposición anidada estocástica de Benders para su resolución. Se han analizado y desarrollado diversas técnicas y mejoras en la resolución de los subproblemas y en la implantación del algoritmo de descomposición anidada para reducir su tiempo de ejecución. El modelo ha sido implantado en el lenguaje de modelado GAMS y utiliza el optimizador CPLEX para la resolución de los problemas de optimización. Se presenta un caso de estudio de muy gran tamaño basado en el sistema eléctrico peninsular español.

Entre los desarrollos futuros previstos está el modelado de las curvas de rendimiento de las centrales hidráulicas mediante poligonales dando lugar a un problema binario mixto estocástico resoluble mediante la misma metodología. Para que el tiempo de resolución sea razonable se necesitan mejoras adicionales en el algoritmo de descomposición.

Referencias

- [1] Benders, J.F. "Partitioning Procedures for Solving Mixed-Variable Programming Problems" *Numerische Mathematik*. Vol 4, pp 238-252. 1962.
- [2] Brooke, A., Kendrick, D. and Meeraus, A. *GAMS Release 2.25 A User's Guide*. GAMS Development Corporation, 1996.
- [3] Jacobs, J, Freeman, G., Grygier, J., Morton, D. P., Schultz, G., Staschus, K. and Stedinger, J. "SOCRATES: A System for Scheduling Hydroelectric Generation under Uncertainty" *Annals of Operations Research*. Vol 59, pp 99-133. 1995.
- [4] Morton, D. P "An Enhanced Decomposition Algorithm for Multistage Stochastic Hydroelectric Scheduling" *Annals of Operations Research*. Vol 64, pp 211-235. 1996.
- [5] Pereira, M.V.F. and Pinto, L.M.V.G. "Multi-Stage Stochastic Optimization Applied to Energy Planning" *Mathematical Programming*. Vol 52, pp 359-375. August 1991.