



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

Tratamiento estadístico de los caudales afluentes en los modelos de planificación del CDEC-SIC.

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL

ELECTRICISTA

MARCELA ALEJANDRA ARREDONDO CÁRDENAS

PROFESOR GUÍA

DENINSON FUENTES DEL CAMPO

MIEMBROS DE LA COMISIÓN

OSCAR MOYA ARAVENA

RODRIGO ESPINOZA VILCHES

SANTIAGO DE CHILE

ENERO 2010

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

Tratamiento estadístico de los caudales afluentes en los modelos de planificación del CDEC-SIC.

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL ELECTRICISTA

MARCELA ALEJANDRA ARREDONDO CARDENAS

COMISIÓN EXAMINADORA	NOTA(n°)	CALIFICACIONES (Letras)	FIRMA
PROFESOR GUÍA			
SR. DENINSON FUENTES DEL C.	:
PROFESOR CO-GUÍA			
SR. OSCAR MOYA ARAVENA	:
PROFESOR INTEGRANTE			
SR. RODRIGO ESPINOZA VILCHES	:
NOTA FINAL EXAMEN DE TÍTULO	:

SANTIAGO DE CHILE

ENERO 2010

“Tratamiento estadístico de los caudales afluentes en los modelos de planificación del CDEC-SIC”

La normativa contenida en el Decreto 291/2008 impone nuevas condiciones en la representación del sistema para los procesos de Planificación de la Operación del CDEC-SIC. Estas nuevas condiciones en la representación incrementarán el esfuerzo computacional asociado al algoritmo de Programación Dinámica Dual Estocástica que soluciona el problema de coordinación hidrotérmica, con el consecuente aumento en los tiempos de ejecución del software asociado, lo que puede comprometer el cumplimiento de los plazos con los que cuenta el Departamento de Planificación de la Operación del CDEC-SIC.

Motivado por lo anterior se propone como objetivo de este trabajo de título, la evaluación de nuevas alternativas en el tratamiento de las series hidrológicas que alimentan el modelo de Programación de Largo Plazo (PLP) para obtener mejoras en la precisión de la solución y los tiempos de ejecución. Lo que se busca es reducir el número de hidrologías muestreadas durante la fase de optimización del modelo (aperturas). Para esto, en una primera instancia, se consideran: el Análisis de Componentes Principales, la aplicación de Modelos Temporales para la obtención de Series Sintéticas y el Agrupamiento de Hidrologías en Clusters; eligiendo esta última opción. Se propone la selección de k aperturas a partir de la estadística actual de 48 años hidrológicos, utilizando como referencia los centroides resultantes de la agrupación de esta estadística en k grupos o clusters. Con esto se espera incluir en las k aperturas escogidas, condiciones hidrológicas diferentes entre sí, a fin de mantener la variabilidad de la muestra original.

El trabajo comienza con una etapa exploratoria donde se evalúa la reducción desde 48 a 24, 12, 6 y 3 aperturas. Se constata un descenso del tiempo de ejecución al disminuir el número de aperturas y se obtienen resultados acordes con los correspondientes a la metodología actual, analizándose para estos efectos, las series de: costos marginales en la barra Quillota 220 kV, costos térmicos por etapa y valor del agua en los embalses del SIC. A partir de los buenos resultados de la primera etapa, se analizan opciones intermedias entre 6 y 12 aperturas, las que resultan satisfactorias en tiempos de ejecución y precisión de la solución. Finalmente, se establecen criterios de selección del número de clusters a utilizar y, por ende, del número de aperturas; éstos son: el tiempo de ejecución del programa y el índice de validación de clusters “silhouette”. Del primero y, a fin de garantizar que la muestra sea insesgada, se obtiene que la replicación de aperturas no incrementa el tiempo de ejecución con respecto al caso que no las replica. El segundo criterio permite distinguir qué particiones realizadas no corresponden a una estructura natural de grupos de la muestra original, a partir del análisis del valor de “silhouette” promedio y de sus gráficas asociadas.

El trabajo desarrollado determina la selección de un menor número de aperturas, manteniendo la precisión de los resultados y disminuyendo los tiempos de ejecución. Esta memoria selecciona las aperturas a partir de la energía afluente global del sistema, por lo que queda abierta la posibilidad de definir aperturas para cada cuenca o embalse separadamente, lo cual requeriría la intervención del código del Modelo PLP. En un sentido más amplio, poder modificar la programación del Modelo PLP permitiría evaluar, no sólo las aperturas de la fase de optimización, sino otros aspectos de la modelación del problema con miras a la reducción de los tiempos de ejecución.

A mis padres, Laura y Carlos, y a Patricio.

Agradecimientos

En primer lugar quisiera agradecer a mi familia: a mis padres, porque esto es parte de un camino que trazaron hace muchos años para mí y mis hermanas, con su esfuerzo y sacrificio, y a mis hermanas por su preocupación y apoyo.

También, quiero agradecer a Patricio porque siempre me ha dado su apoyo incondicional; alentándome, acompañándome y dándome fuerzas para seguir adelante.

De igual manera, quiero hacer mención a mis amigos y amigas, tanto a aquellos que conocí en Beauchef, como aquellas con las que mantenemos una amistad desde los tiempos del colegio. Me siento contenta y agradecida de haber compartido con ustedes en el transcurso de estos años.

Es importante para mí reconocer la importancia que tuvo en mi formación haber participado en la Escuela de Verano, pues siendo alumna aprendí a aprender y definí mi vocación, y, años después, como docente pude enseñar y aprender. De igual manera, ser parte del grupo de Difusión Colegios, me permitió desarrollar habilidades que no se aprenden en una cátedra, pero que son fundamentales en lo laboral y personal.

Quisiera agradecer al Centro de Despacho Económico de Carga del Sistema Interconectado Central por el financiamiento de este trabajo. También agradezco a su personal por haberme acogido durante el desarrollo de esta memoria, así como por su interés y apoyo.

También extendo mis agradecimientos a Rodrigo Espinoza Vilches por su valiosa contribución al desarrollo de esta memoria.

Finalmente, quisiera agradecer especialmente a Deninson Fuentes del Campo, mi profesor guía. Gracias por su buena disposición para corregir los borradores de este trabajo, por su generosidad al compartir sus conocimientos y la confianza depositada en mí.

CONTENIDO

1	Introducción.....	1
1.1	Motivación.....	1
1.2	Objetivos	2
1.2.1	Objetivo General.....	2
1.2.2	Objetivos Específicos.....	2
1.3	Estructura del Documento.....	2
2	Antecedentes.....	3
2.1	El Problema de Coordinación Hidrotérmica.....	3
2.1.1	Programación Dinámica	4
2.1.2	Programación Dinámica Estocástica	5
2.1.3	Programación Dinámica Dual Estocástica	8
2.1.3.1	Generalidades.....	8
2.1.3.2	Modelos.....	10
2.2	Aleatoriedad de los Afluentes de las Cuencas.	11
2.2.1	Naturaleza del Comportamiento	11
2.2.2	Características Estadísticas de las Series Hidrológicas.....	12
2.2.3	Series Hidrológicas Utilizadas en el CDEC-SIC.....	12
2.2.3.1	Matriz de energía afluente.....	13
3	Selección de Aperturas	15
3.1	Alternativas de Caudales Afluentes.....	15
3.1.1	Análisis de Componentes Principales.....	15
3.1.2	Series Sintéticas.....	15
3.1.3	Método de Análisis de Clusters	16
3.1.4	Comparación y Selección.....	16
3.2	Agrupamiento de Hidrologías en Clusters	17
3.2.1	Análisis Exploratorio.....	17
3.2.1.1	Metodología.....	17
3.2.1.2	Resultados y Análisis.....	18
3.2.2	Evaluación de Alternativas de Agrupamiento.....	27
3.2.2.1	Metodología.....	27
3.2.2.2	Resultados y Análisis.....	28
3.2.3	Criterios de Selección	32
3.2.3.1	Tiempos de Ejecución.....	32
3.2.3.2	Silhouette	33
3.2.3.3	Síntesis.....	36
4	Conclusiones	37

Referencias.....	39
ANEXO A: Elementos Matemáticos	A1
Clusters	A1
Antecedentes generales.....	A1
Distancia Euclideana.....	A2
Algoritmo K-Means	A2
Algoritmo convencional de k-means.....	A3
ANEXO B: Descripción de Etapas	B1
ANEXO C: Resultados PLP	C1
Comparación de resultados	C1
Problemas lineales en CPLEX.....	C15
ANEXO D: Resultados de Agrupación de Clusters	D1

Índice de Figuras

Figura 2-1: Árbol de decisión para Sistemas Hidrotérmicos con centrales de embalse.....	3
Figura 2-2: Funciones de Costo Inmediato (FCI) y Costo Futuro (FCF) de acuerdo al volumen final de agua en el embalse.	4
Figura 2-3: Definición de las variables de estado	6
Figura 2-4: Cálculo de la estrategia óptima en la última etapa $t=T$	6
Figura 2-5: Función de costo futuro esperado para etapa $T-1$	6
Figura 2-6: Minimización de costos en etapa $T-1$	7
Figura 2-7: Cálculo del primer segmento de la aproximación lineal de la FCF con PDDE	8
Figura 2-8: Cálculo de la FCF como aproximación lineal por partes para la etapa $T-1$	9
Figura 2-9: Esquema de aperturas para simulación i -ésima	11
Figura 2-10: Ciclo Hidrológico.....	12
Figura 3-1: Esquema de trabajo propuesto	18
Figura 3-2: Gráficos de Costos Marginales para años: Húmedo, medio y seco.....	26
Figura 3-3: Elementos involucrados en el cálculo de $s(i)$	33
Figura 3-4 Gráfico de silhouette entre 6 y 9 clusters.....	34
Figura 3-5: Gráfico de silhouette Gráfico de silhouette entre 6 y 9 clusters.....	35

Índice de Gráficos

Gráfico 2-1: Probabilidad de Excedencia.....	14
Gráfico 3-1: Tiempo de Ejecución PLP según número de aperturas	22
Gráfico 3-2: Costos Marginales en la barra Quillota 220, valor esperado para el caso 17/04/2008.....	23
Gráfico 3-3: Costos térmicos por etapas, valor esperado para el caso del 17/04/2009	23
Gráfico 3-4: Tiempo de Ejecución PLP según número de aperturas	30
Gráfico 3-5: Costos marginales en la Barra Quillota220, Valor esperado para el caso del 17/04/2009.....	31
Gráfico 3-6: Costos térmicos por etapas, valor esperado para el caso del 17/04/2009	31
Gráfico 3-7: Tiempos de Ejecución según número de aperturas(ojo con el Label de los puntos Azules).....	32
Gráfico 3-8: Tiempos de Ejecución según número de clusters	32

1 INTRODUCCIÓN

1.1 MOTIVACIÓN

El problema de coordinación hidrotérmica corresponde a la optimización simultánea del abastecimiento de la demanda de una red eléctrica y el uso del recurso hidráulico en las cuencas y embalses del sistema. Dependiendo del horizonte de planificación, el problema adquiere distintas características, pasando desde un problema determinístico cuando se soluciona para el corto plazo, a uno estocástico en el largo plazo, donde la demanda y las hidrologías futuras constituyen variables aleatorias. Para abordar este problema se han desarrollado diferentes alternativas, en particular, durante la década de los noventa se consolidó el uso de la denominada Programación Dinámica Dual Estocástica (PDDE), la cual se basa en el algoritmo de descomposición anidada de Benders, que permite resolver en forma más eficiente problemas de optimización de gran tamaño que tienen una estructura de bloques en sus restricciones. [1]

En el país, el CDEC-SIC hace uso de esta herramienta matemática a través del modelo de Programación de Largo Plazo (PLP), desarrollado por Colbún S.A.

La Programación de Mediano y Largo Plazo en el CDEC-SIC “derivará de estudios de planificación de la operación del sistema eléctrico que, preservando la seguridad de servicio instantánea global del sistema, lleven a minimizar su costo total actualizado de operación y de falla para un horizonte de 5 años” según la normativa impuesta por el Decreto N°291/2008 [2]. Como consecuencia de la entrada en vigencia de este Decreto aparecen elementos de modelación adicionales a los procesos de Planificación de la Operación como son: la discretización del horizonte de planificación en un mayor número de etapas, la modelación temporal de la demanda en al menos tres bloques para cada una de las etapas y la modelación del sistema de transmisión hasta el nivel de 66 [kV], entre otros. La aplicación de esta nueva modelación del sistema y sus elementos al PLP tendrá una importante repercusión en los tiempos de ejecución del programa, previéndose un incremento en estos tiempos que puede resultar perjudicial para el cumplimiento de los plazos que dispone la Dirección de Operación. Es por esta razón que como trabajo de memoria se propone la exploración de nuevas alternativas al tratamiento hidrológico tradicional en la Programación Dinámica Dual Estocástica que resuelve el problema de coordinación hidrotérmica en el CDEC-SIC, para obtener mejoras en la precisión de la solución y tiempos de ejecución.

Este trabajo propone reducir el número de aperturas que se utilizan en la fase de optimización del PLP, seleccionando algunas de las hidrologías contenidas en la estadística histórica, a partir de la agrupación de éstas en clusters. Así, agrupando en clusters y seleccionando, en una primera instancia, una hidrología por cada cluster, se reduce el tamaño de la muestra, cuidando mantener en ella los elementos que explican la mayor parte de la variabilidad de la muestra.

1.2 OBJETIVOS

1.2.1 OBJETIVO GENERAL

El objetivo principal del trabajo es evaluar el desempeño en el uso de la metodología de programación dinámica dual estocástica para la solución del problema de coordinación hidrotérmica, mediante ajustes en el tratamiento de las series hidrológicas que alimentan al modelo.

1.2.2 OBJETIVOS ESPECÍFICOS

Los objetivos específicos se detallan a continuación:

- Revisar el estado del arte de las metodologías que resuelven el problema de coordinación hidrotérmica, particularmente lo referente a la PDDE.
- Comprender la metodología utilizada en el CDEC-SIC en sus procesos de planificación de la operación con respecto al tratamiento hidrológico (dimensión normativa y de modelo).
- Estudiar las series hidrológicas utilizadas por el CDEC-SIC en sus modelos.
- Evaluar la calidad de las eventuales modificaciones al tratamiento hidrológico en la PDDE en cuanto a la precisión de la solución y tiempos de ejecución.
- Proponer la/las mejores estrategias de aplicación de la estadística de caudales afluentes, durante la fase de optimización, en los modelos de planificación de la operación.

1.3 ESTRUCTURA DEL DOCUMENTO

El documento está estructurado como sigue:

- El Capítulo 2 recoge las características que tiene el problema de coordinación hidrotérmica haciendo especial énfasis en las herramientas matemáticas que lo resuelven como son: la Programación Dinámica Estocástica y la Programación Dinámica Dual Estocástica. Estos temas se abordan de una manera descriptiva, así como se presentan modelos basados en estas metodologías en Chile y el mundo. Por otra parte, en este capítulo se incluyen los antecedentes necesarios para entender la estructura de las estadísticas hidrológicas y el manejo que se realiza de los caudales afluentes a las distintas cuencas del sistema, por parte del CDEC-SIC.
- En el Capítulo 3 se presenta la metodología de selección de aperturas. Este capítulo se ha organizado de una forma secuencial, mostrando los pasos que se siguieron desde una propuesta exploratoria que permitió tomar las primeras decisiones, para continuar con las opciones de agrupamiento en clusters y, finalmente, presentar los criterios de selección de aperturas.
- Por último, en el capítulo 4 se encuentran las conclusiones más importantes del trabajo desarrollado.

2 ANTECEDENTES

En este capítulo se incluyen los principales antecedentes para contextualizar el trabajo de título a desarrollar, en la primera parte se presentan las características generales de la coordinación hidrotérmica en sistemas eléctricos, junto con las distintas herramientas de modelación con las que se ha abordado este problema. La segunda parte de este capítulo está dedicada a la naturaleza estocástica de los caudales afluentes de las centrales de embalse, analizando la naturaleza de su comportamiento, sus características estadísticas y revisando la estadística de caudales disponible en el CDEC-SIC.

2.1 EL PROBLEMA DE COORDINACIÓN HIDROTÉRMICA.

El despacho de un sistema eléctrico es un problema que presenta distintas características de acuerdo al horizonte de planificación con el que se esté trabajando. Sin embargo, una vez que se cuenta con el parque generador del sistema, se realiza la planificación de la operación de estas unidades generadoras para determinar políticas que permitan su funcionamiento de la forma más económica y segura posible.

De acuerdo a las diferentes tecnologías disponibles en el parque generador, el problema de despacho de las unidades tendrá distintas características. Por ejemplo, en un sistema principalmente térmico la política de operación será la “orden de mérito”, en donde las unidades generadoras entran al despacho en forma creciente según sus costos variables hasta satisfacer la demanda del período, pudiendo repetir esta política en etapas futuras, mientras el combustible (gas, carbón, petróleo, etc.) se encuentre disponible. Sin embargo, para la operación de un sistema eléctrico con centrales hidroeléctricas de embalse que permitan regulación interanual, el problema se acopla en el tiempo, ya que las decisiones de operación que se tomen en el presente afectarán el estado del sistema en el futuro.

Para ilustrar el acoplamiento en el tiempo de las políticas de operación de una central de embalse, el problema puede representarse como el árbol de decisión de la Figura 2-1, de acuerdo al esquema seguido por Pereira *et al.* en [3].

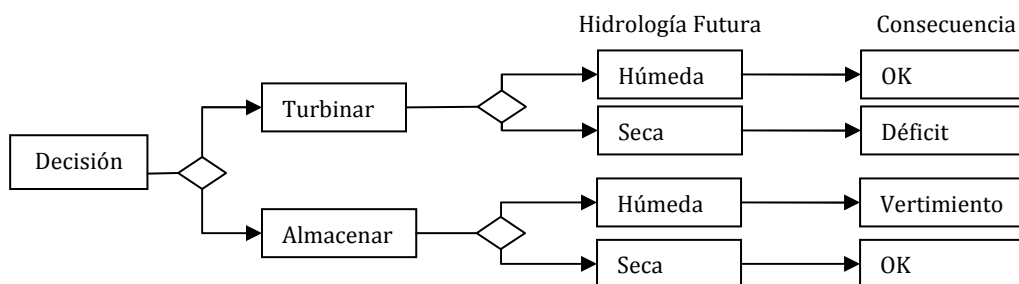


FIGURA 2-1: ÁRBOL DE DECISIÓN PARA SISTEMAS HIDROTÉRMICOS CON CENTRALES DE EMBALSE

Como puede verse en la figura, el operador tiene la opción de turbinar el agua del embalse hoy o almacenarla para su uso en un futuro. Al elegir la primera opción, se minimizan los costos térmicos de la etapa actual, a la vez que se disminuyen las reservas de agua en el embalse; por lo que, de producirse en el futuro una condición hidrológica húmeda, la operación del sistema resultará eficiente. Sin embargo, si a futuro acontece una condición hidrológica seca, los costos que afrontará el sistema serán mayores, pues no se cuenta con las reservas de agua necesarias y se tendrá que generar con centrales térmicas. En la situación de sequía extrema, podría haber racionamiento.

Por otra parte, si se decide almacenar agua en el embalse, incurriendo en mayores costos térmicos en la etapa actual y se presenta una hidrología húmeda, se producirá vertimiento en el embalse, lo que representa una pérdida de energía generada y, por ende, un incremento en los costos de operación. Finalmente, si a futuro hay una hidrología seca, el agua almacenada evitará el uso de generación térmica a alto costo o racionamiento.

De esta forma, para determinar la operación óptima de un sistema hidrotérmico se debe minimizar la suma de los costos de operación térmicos de la etapa actual y de etapas futuras, para cada una de las etapas definidas a lo largo del horizonte de planificación. La dependencia de estos costos con respecto al volumen final de agua embalsada se muestra en la Figura 2-2.

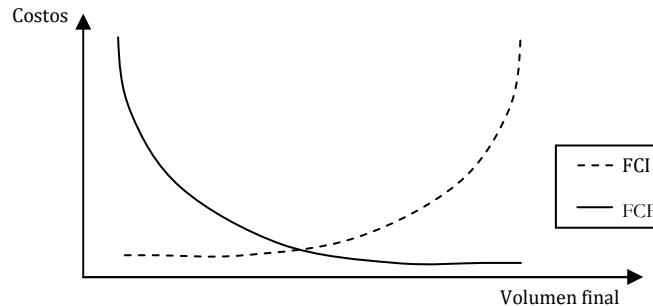


FIGURA 2-2: FUNCIONES DE COSTO INMEDIATO (FCI) Y COSTO FUTURO (FCF) DE ACUERDO AL VOLUMEN FINAL DE AGUA EN EL EMBALSE.

La función de costo inmediato (FCI) corresponde a los costos térmicos del sistema necesarios para satisfacer la demanda de la etapa actual; estos costos se reducen al aumentar la generación hidroeléctrica producida con el agua embalsada, lo que disminuye el volumen final embalsado. Por el contrario, los costos futuros (FCF) aumentan al turbinar el agua del embalse, pues éstos representan la generación térmica esperada para las etapas siguientes, la que se ve incrementada a medida que disminuyen las reservas de agua en el embalse.

Como existen muchos parámetros del sistema que son inciertos en el futuro, la función de costo futuro tendrá un carácter estocástico, siendo el parámetro más relevante la incertidumbre hidrológica. En tal sentido, el operador del sistema debe definir un criterio para enfrentar el problema estocástico, y este es, típicamente, la minimización de los costos futuros esperados. Por ello, en presencia de incertidumbre, la FCF se transforma en la función de costos futuros esperados [4]

De esta manera, mientras la FCI se obtiene directamente de los costos de las unidades térmicas, la FCF se obtiene a partir de simulaciones de la operación del sistema para distintos valores de almacenamiento inicial y condiciones hidrológicas. Una forma eficiente de obtenerla es a través de la programación dinámica.

2.1.1 PROGRAMACIÓN DINÁMICA

La programación dinámica (PD) es un procedimiento sistemático que permite determinar la combinación de decisiones ligadas en el tiempo que entrega un resultado óptimo. Esta técnica se basa en la aplicación del principio de Bellman, el que impone las siguientes condiciones en el problema:

- El sistema debe ser no hereditario, es decir, el estado del sistema resume toda la historia del mismo.

- La función objetivo debe ser de naturaleza aditiva, lo que permite separar el problema global en etapas que pueden ser analizadas en forma independiente, ligando cada una con el resto del problema a través de los estados inicial y final de la etapa.

La aplicación de la PD en la operación de embalses comienza con la discretización de los niveles de almacenamiento del embalse, para luego asociar a cada cota el costo de la operación futura del sistema. Con esto, la decisión óptima dentro de cada etapa será la que minimiza el costo de operación en la etapa más el costo futuro (asociado a la cota del embalse al final de la etapa), de modo que el problema de coordinación hidrotérmica para una etapa t se formula como sigue:

$$\min c_t(u_t) + \alpha_{t+1}(v_{t+1})$$

sujeto a

$$v_{t+1} = v_t - u_t - s_t + a_t$$

$$v_{t+1} \leq \bar{v}$$

$$u_t \leq \bar{u}$$

Donde " v_t " representa el almacenamiento, " u_t " las decisiones de generación, " s_t " los vertimientos, " a_t " los afluentes, todos en la etapa t ; mientras que $\alpha_{t+1}(v_{t+1})$ es el costo de la etapa futura ($t+1$) como función del almacenamiento.

Si se conoce la función de valores estratégicos del agua al final del último período del horizonte de análisis, se puede determinar la decisión óptima para cada nivel inicial posible en este último período. Avanzando en sentido inverso del tiempo se determina para cada etapa y nivel de embalse, la decisión óptima y el valor estratégico asociado. Se forma así una malla de decisiones óptimas hasta llegar al primer período en que el nivel inicial es conocido. Recorriendo la malla de decisiones óptimas desde el punto inicial conocido, ahora en el sentido del tiempo, se encuentra la secuencia de decisiones óptimas y la trayectoria de la cota del embalse.

2.1.2 PROGRAMACIÓN DINÁMICA ESTOCÁSTICA

El planteamiento general de la PD señalado anteriormente no toma en cuenta la naturaleza estocástica de las hidrologías que determinarán los caudales afluentes a las centrales de embalse, por lo que la programación dinámica estocástica (en adelante, PDE) aparece como una alternativa de solución para la coordinación hidrotérmica ante condiciones hidrológicas futuras inciertas.

El algoritmo de PDE tiene dos etapas: una de optimización y una de simulación. La primera etapa utiliza una *recursión backward*, la que puede ser descrita, en términos generales, por los siguientes pasos, de acuerdo al esquema propuesto por Pereira *et al.* en [3]:

- a) Se discretiza el recorrido de la variable de estado para cada etapa de tiempo t , por ejemplo, si la variable de estado son los niveles de almacenamiento en un embalse, éstos pueden ser discretizados en 100%, 90%, etc. hasta 0%. Se supone conocido el estado inicial del problema.

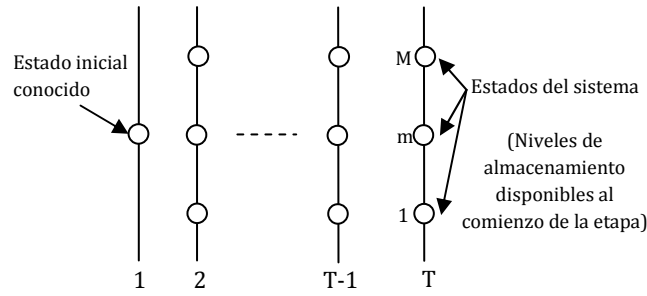


FIGURA 2-3: DEFINICIÓN DE LAS VARIABLES DE ESTADO

- b) El algoritmo comienza en la última etapa de evaluación, resolviendo el problema de despacho considerando el primer nivel de almacenamiento definido en a) para todos los K escenarios de caudales posibles.

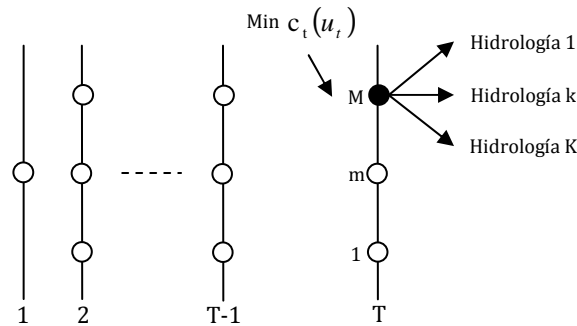


FIGURA 2-4: CÁLCULO DE LA ESTRATEGIA ÓPTIMA EN LA ÚLTIMA ETAPA T=T

- c) Si se supone que todos los escenarios de caudal son equiprobables, el valor esperado del costo de operación asociado al nivel de almacenamiento M, será el promedio de los K subproblemas resueltos. Así se obtiene el primer punto de la función de costo futuro esperado (FCFE) para la etapa anterior, tras lo cual se debe repetir el proceso para todos los niveles de almacenamiento del embalse. Al interpolar los puntos obtenidos para cada nivel de almacenamiento se obtiene la FCFE $\alpha_T(v_T)$ para la etapa T-1.

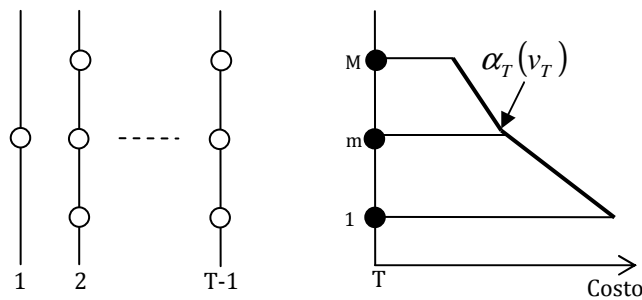


FIGURA 2-5: FUNCIÓN DE COSTO FUTURO ESPERADO PARA ETAPA T-1

- d) El proceso anterior es repetido para todas las etapas desde la T-1 hacia la primera etapa. En cada una de estas etapas el objetivo será minimizar el costo de operación de la etapa más el costo futuro, dado por la FCFE calculada previamente.

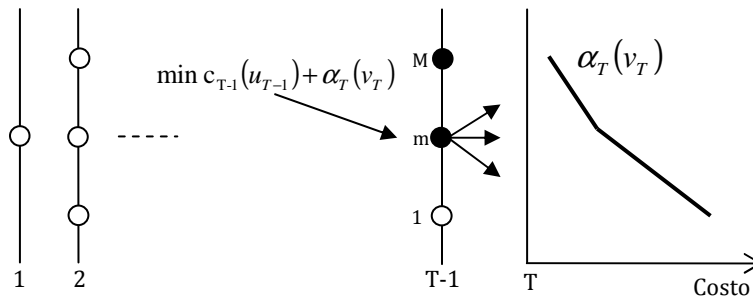


FIGURA 2-6: MINIMIZACIÓN DE COSTOS EN ETAPA T-1

La fase de simulación realiza una *iteración forward* en la que se obtienen las cotas óptimas del embalse en cada etapa, partiendo desde la primera etapa, donde el estado es conocido, hacia la última.

La principal limitación de este método viene dada por la necesidad de enumerar todas las combinaciones posibles de las variables de estado, lo cual hace explotar combinatorialmente la representación de estados, ya sea al aumentar el número de embalses y/o al afinar la discretización de estados.

Dentro de los modelos de planificación utilizados a lo largo de la historia en el Sistema Eléctrico Chileno, los modelos GOL (Gestión Óptima del Laja, 1974) y OMSIC (Operación Mensual del SIC) se han basado en la aplicación de la PDE.

El modelo GOL fue desarrollado como una herramienta computacional para estimar la demanda de carbón del sector eléctrico y fue utilizado por Endesa para la programación de la operación de las centrales de la cuenca del Laja, para un horizonte de planificación de hasta 15 años en etapas trimestrales. [5]

Este modelo busca minimizar el costo esperado de operación de las centrales térmicas, a través del uso óptimo del agua embalsada en el Laja y despachando a las unidades térmicas de acuerdo a la “lista de mérito”. Para esto se representa el SIC en forma uninodal, con demanda trimestral determinística y en donde las centrales del Laja y las centrales térmicas se representan individualmente, mientras que las centrales hidroeléctricas fuera de la cuenca del Laja se agrupan en una única central de pasada.

El tratamiento dado a la aleatoriedad hidrológica puede encontrarse en detalle en [6]. A continuación se presentan las características más importantes:

- Es un hecho que existe independencia estadística entre los caudales afluentes de cada año. Es decir, la ocurrencia de una hidrología durante un año no determina ni condiciona la hidrología del año siguiente. Al no ser posible prever lo que ocurrirá en los años siguientes, el valor estratégico asociado a una cota a fines de cada año hidrológico (abril a marzo) corresponde al valor esperado respecto a las hidrologías futuras posibles.
- Para un año, la decisión de extracción del embalse supone conocida la hidrología del año (esquema azar-decisión). Para cada hidrología anual, supuesta conocida, el sistema se opera en forma óptima y el valor estratégico a comienzos del año, para cada cota, es calculado promediando los valores estratégicos de todas las hidrologías posibles.
- El año se ha dividido en subetapas trimestrales con el fin de representar la estacionalidad de la demanda, los aportes hidrológicos y la disponibilidad de las centrales que varían a lo largo del año. Los aportes hidráulicos en cada trimestre se suponen perfectamente

dependientes entre sí, es decir, la operación óptima a través de los cuatro trimestres se realiza como una programación dinámica determinística para cada hidrología.

Por su parte, el modelo OMSIC fue utilizado por el CDEC-SIC para la programación de la operación de mediano y largo plazo, entre los años 1991 y 2004. Este modelo fue desarrollado a partir del modelo GOL, por lo que la metodología de solución del problema de la gestión óptima del SIC es la PDE, planteando cambios en la duración de las etapas, en el horizonte de planificación y en el tratamiento de la hidrología. De esta manera, el OMSIC trabaja con etapas mensuales salvo por el primer mes que se divide en períodos semanales; con un horizonte de planificación de uno o dos años y considerando que las hidrologías de invierno son estadísticamente independientes, mientras que las de verano son estadísticamente dependientes.

2.1.3 PROGRAMACIÓN DINÁMICA DUAL ESTOCÁSTICA

2.1.3.1 Generalidades

Al igual que la PDE, la Programación Dinámica Dual Estocástica (PDDE) permite determinar la política de operación óptima de las centrales de embalse y cada iteración cuenta con una etapa *forward* y una *backward*. Es en esta última donde se diferencian, ya que la PDDE utiliza información del problema dual para construir hiperplanos que se aproximan inferiormente a la FCF, extrapolando esta función a partir de la evaluación para algunos niveles de almacenamiento. Lo que permite resolver problemas de despacho estocástico con un gran número de embalses, con un esfuerzo computacional razonable.

El esquema usado en PDDE se basa en que la FCF puede ser representada por una función lineal por partes, es decir, no es necesario generar una tabla de valores a ser interpolados para aproximar la FCF. Es más, de acuerdo a la Figura 2-7, la pendiente de la FCF en un punto puede ser obtenida en forma analítica a partir del problema de despacho en una etapa.

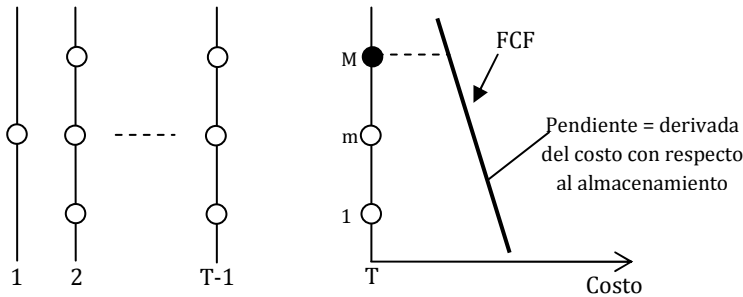


FIGURA 2-7: CÁLCULO DEL PRIMER SEGMENTO DE LA APROXIMACIÓN LINEAL DE LA FCF CON PDDE

Para el caso de este último período, el problema de minimización de costos en la etapa es:

$$\begin{aligned}
 z_T &= \min c_T(u_T) && \text{Multiplicadores} \\
 \text{sujeto a} &&& \\
 v_{t+1} &= v_t - u_t - s_t + a_t && \pi_h \\
 v_{t+1} &\leq \bar{v} && \pi_v \\
 u_t &\leq \bar{u} && \pi_u
 \end{aligned}$$

En la formulación anterior, la función de costo futuro de esta última etapa se establece como cero. Además, asociado a cada restricción del problema en el punto óptimo, aparece el multiplicador simplex correspondiente. Estos multiplicadores corresponden a las derivadas de la solución óptima con respecto a las variables de decisión, es decir, del costo de operación con respecto a una perturbación en el lado derecho de cada restricción. Dentro de estos multiplicadores, el asociado a la ecuación de balance hidráulico, π_h , es la derivada de z_T con respecto a una variación en almacenamiento inicial v_T :

$$\pi_h = \frac{\partial z_T}{\partial v_T}$$

Como se describe en la Figura 2-7, la expresión anterior es la pendiente de la FCF para la etapa T-1, cabiendo también su interpretación como una expansión en serie de primer orden de la FCF en torno al almacenamiento inicial v_T .

La Figura 2-8 muestra el cálculo de los costos de operación y las pendientes de la FCF para cada estado en la etapa T. La FCF $\alpha_T(v_T)$ para la etapa T-1 es la superficie lineal que resulta al escoger los segmentos con los mayores costos en cada estado.

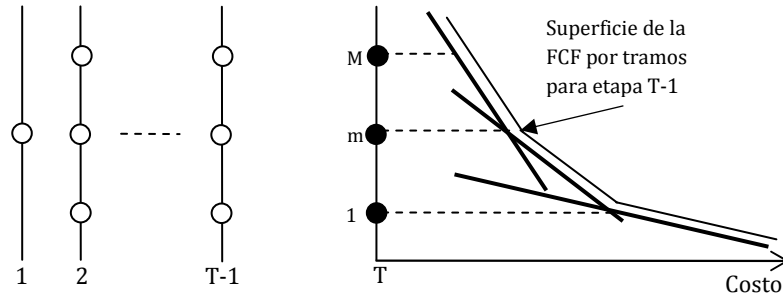


FIGURA 2-8: CÁLCULO DE LA FCF COMO APROXIMACIÓN LINEAL POR PARTES PARA LA ETAPA T-1

Finalmente el problema de despacho hidrotérmico para la etapa T-1 se presenta como un problema de programación lineal.

$$\begin{aligned} \alpha_{T-1}(v_{T-1}) &= \min c_{T-1}(u_{T-1}) + \alpha_T(v_T) \\ \text{sujeto a} \\ v_T &= v_{T-1} - u_{T-1} - s_{T-1} + a_{T-1} \\ v_T &\leq \bar{v} \\ u_{T-1} &\leq \bar{u} \\ \alpha_T &\geq \varphi_T^n \cdot v_T + \delta_T^n \quad n = 1, \dots, N \end{aligned}$$

La FCF es representada por la variable escalar α_T y las N restricciones lineales $\varphi_T^n \cdot v_T + \delta_T^n$, donde N es el número de segmentos lineales.

Es conveniente notar las diferencias que tienen la PDE y la PDDE. En el esquema tradicional, la FCF se obtiene al interpolar los distintos estados de almacenamiento del embalse (variable discreta), lo que obliga a la evaluación de todos los estados. Mientras que en la PDDE, los segmentos lineales son utilizados para extrapolar los valores de la FCF, de manera que no es necesario usar todas las combinaciones de puntos para obtener la FCF para todos los estados. [7] El algoritmo mediante el cual se escogen las aproximaciones tangentes se denomina descomposición de Benders y las rectas reciben

el nombre de planos minorantes o cortes de Benders. En el algoritmo de Benders, los cortes se agregan uno a uno por iteración escogiendo convenientemente, mediante simulaciones de la operación del sistema, los estados donde se calcula la función hasta alcanzar la convergencia.

Estas diferencias han permitido que modelos basados en la metodología de PDDE, aborden el problema de coordinación hidrotérmica en sistemas multiembalses y multinodales en tiempos de ejecución razonables. Algunos de estos modelos son SDDP (Brasil) y PLP (Chile).

2.1.3.2 Modelos

El modelo SDDP (Stochastic Dual Dynamic Programming) desarrollado por PSR Inc. (Power Systems Research), está basado en la programación dinámica dual estocástica, con búsqueda del óptimo mediante programación lineal. Este modelo entrega una representación continua de la FCF en cada etapa, para el problema multinodal y multiembalse, haciendo una representación lineal de las pérdidas en el sistema de transmisión. Su horizonte de planificación es variable y suele ser de 10 años, más dos años de relleno. Las etapas de optimización pueden ser semanales o mensuales.

En el sistema chileno, el modelo PLP (Programación de Largo Plazo) es utilizado por el CDEC-SIC (Centro de Despacho Económico de Carga del Sistema Interconectado Central) dentro de sus procesos de planificación de la operación. El PLP fue desarrollado por Colbún S.A., basándose en la PDDE de forma análoga al SDDP; abordando el problema de coordinación hidrotérmica del SIC en forma multimodal y multiembalse, representando las pérdidas de transmisión por tramos lineales e incluyendo la modelación de los convenios de riego del Río Maule y del Río Laja.

Históricamente, el tratamiento de la hidrología que realiza el modelo PLP estuvo regulado por el Artículo 190 del D.S. N°327/1997 de acuerdo a lo siguiente “...la estadística para los gastos afluentes en centrales hidroeléctricas deberá ser representativa de las respectivas centrales y considerar, como mínimo, una muestra de 40 años hidrológicas. Esta estadística deberá ser actualizada periódicamente, de modo que siempre incluya los datos correspondientes al año hidrológico anterior al que precede a aquel en que se esté efectuando la programación de la operación”.

Se adoptó que los 40 escenarios hidrológicos históricos constituyen una estadística *representativa*, siguiendo el mismo registro observado en su oportunidad, donde los años simulados se encuentran separados interanualmente en la forma de años hidrológicos (abril-marzo)

Adicionalmente, el año es dividido en dos períodos debido a la naturaleza nivopluvial de los regímenes de caudales de las cuencas más importantes del Sistema Interconectado Central (SIC), los regímenes de este tipo de cuencas se caracterizan por dos temporadas claramente marcadas, una temporada de invierno y una temporada de deshielo. La calidad de la primera temporada determinada la calidad de la segunda cuyos caudales son alimentados fundamentalmente por los derretimientos de nieve acumulada durante el período invernal. Esta separación del año en dos períodos, conocidos como independencia hidrológica e incertidumbre reducida, desde abril a septiembre y de octubre a marzo respectivamente, determina que se incorporen aspectos particulares, tanto en la fase de simulación como optimización de los modelos de embalse, a excepción de aquellas centrales que siguen exclusivamente un régimen pluvial. En estas últimas, los escurrimientos durante el período octubre a marzo están determinados por las precipitaciones que se puedan registrar al interior del mismo período, por lo que están sujetos a una mayor variabilidad hidrológica. [8]

Conforme a la metodología de la PDDE, el modelo PLP calcula la función de costos futuros a través de la iteración “hacia atrás” definiendo aperturas para cada estado y cada etapa, para luego simular la operación del sistema “hacia delante”. Para todo este proceso las n hidrologías disponibles

son numeradas desde 1 a n según orden cronológico, este número es utilizado como índice para identificar cada año hidrológico durante la selección de aperturas o simulaciones, definidas por el usuario en los archivos de entrada del PLP¹.

Las aperturas asociadas a la recursión *backward* se definen como un muestreo de escenarios hidrológicos que representan la incertidumbre hidrológica futura que deberá enfrentar el planificador al resolver el problema desde el presente hacia el futuro en el marco de la etapa de simulación. Para las centrales que exhiben un régimen nivo-pluvial, se exige que durante el período de deshielo sigan sólo una apertura y que adicionalmente, ésta coincida con la simulación que definió el estado en torno al cual se está estimando la aproximación de la FCF.

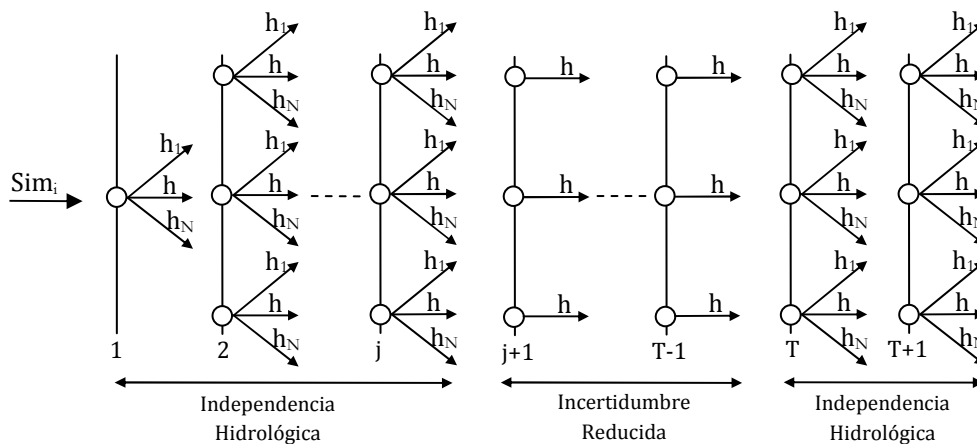


FIGURA 2-9: ESQUEMA DE APERTURAS PARA SIMULACIÓN I-ÉSIMA.

En la Figura 2-9 se ejemplifica el esquema de las aperturas para la i -ésima simulación. De acuerdo a lo anterior, durante la recursión *backward* se definen N aperturas por cada estado durante los períodos de independencia hidrológica, mientras que en el período de incertidumbre reducida la función de costos futuro se calcula con una única apertura (que es la hidrología correspondiente a la de la simulación). De esta manera, en la recursión *forward* el problema de minimización de los costos presentes y futuros, considerará el promedio de los costos futuros obtenidos para cada apertura en el período de independencia hidrológica (es decir, se consideran todas las hidrologías como equiprobables para obtener el costo futuro esperado) y, durante el período de incertidumbre reducida, la función de costo futuro es obtenida directamente del costo futuro calculado utilizando la hidrología i -ésima como única apertura.

2.2 ALEATORIEDAD DE LOS AFLUENTES DE LAS CUENCAS.

2.2.1 NATURALEZA DEL COMPORTAMIENTO

La naturaleza del comportamiento de los afluentes a las cuencas está determinada por el ciclo hidrológico. Se entiende por ciclo hidrológico al complejo sistema dinámico que caracteriza el agua en el planeta, producido por la energía que emana del sol y alcanza la tierra. Como puede observarse en la Figura 2-10, los caudales afluentes a una cuenca están determinados por precipitaciones y deshielos inmersos dentro del ciclo hidrológico, por lo que se encuentran bajo numerosas condiciones estocásticas que determinarán la alta aleatoriedad de estos caudales.

¹El modelo recibe estos parámetros a través de los archivos IPLPIDSIM.DAT, PLPIDAPE.DAT y PLPIDAPE2.DAT. El primer archivo contiene los índices para cada una de las simulaciones, y los siguientes, los índices de las aperturas por simulación para las centrales con régimen nivopluvial y para las centrales con régimen pluvial, respectivamente.

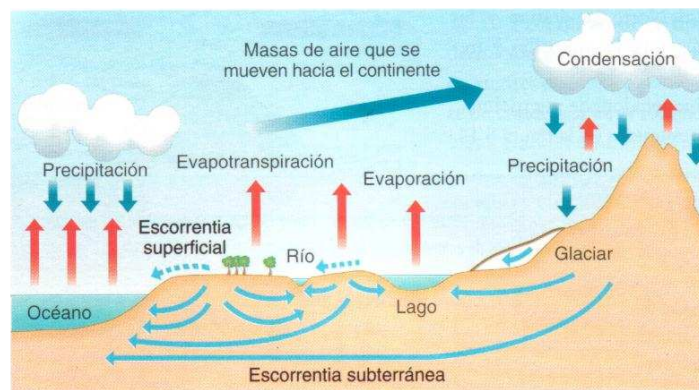


FIGURA 2-10: CICLO HIDROLÓGICO

2.2.2 CARACTERÍSTICAS ESTADÍSTICAS DE LAS SERIES HIDROLÓGICAS

Se presentan las características estadísticas de las series hidrológicas y que las diferencias de otras series temporales. Entre estas características se puede destacar que su valor esperado y desviación estándar son periódicos para intervalos de tiempo menores de un año. Este comportamiento está condicionado por el movimiento astronómico de la Tierra, que determina ciclos periódicos en las series geofísicas. Por otro lado, las series hidrológicas se componen de valores positivos o nulos, por lo que son series no negativas. Otras características a destacar son las que siguen:

1. *No Estacionaridad*: discretizado en intervalos menores a un año, la media, varianza y covarianza de la serie hidrológica varían con respecto al tiempo. Para intervalos mayores que un año la serie de afluentes promedios tiene características estacionarias [9].
2. *Estacionalidad*: da cuenta del comportamiento cíclico de la serie, donde el caudal de los meses de verano está fuertemente correlacionado con las precipitaciones ocurridas en la cordillera en los meses de invierno (régimen nivopluvial).
3. *Dependencia Espacial*: puede observarse correlación espacial entre series hidrológicas, evidenciándose relación entre caudales medidos en distintas posiciones geográficas.
4. *Intermitencia*: al discretizar en escalas de tiempo muy pequeñas, la serie puede alternar entre un estado seco y uno húmedo, de acuerdo a ausencia o presencia de precipitaciones, transformando la serie en una variable binaria. En general, para etapas de duración superior a un mes, la intermitencia del fenómeno es en promedio despreciable [10]

2.2.3 SERIES HIDROLÓGICAS UTILIZADAS EN EL CDEC-SIC.

La estadística de las series hidrológicas disponible en el CDEC-SIC se recoge en “años hidrológicos” que van desde Abril de un año a Marzo del año siguiente, a fin de registrar el período de precipitaciones de invierno, junto con el período de deshielo consecutivo, dada la correlación esperada entre ambos. Estas series se comprenden de valores “semanales”, en donde la duración de cada semana es de 7 u 8 días, a fin de obtener 48 semanas en el año hidrológico.

Las series de información de caudales disponibles provienen de mediciones que se han realizado en las diversas cuencas del país. Los valores considerados son obtenidos fundamentalmente de tres formas [11]:

1. Medición directa de la serie en estaciones fluviométricas ubicadas en el lugar considerado.
2. Obtención de los valores a través de la adición y substracción de flujos medidos en otros lugares de la cuenca, dado que en el lugar de interés no se cuenta con mediciones directas. De esta manera se espera completar la información necesaria, en base a relaciones de continuidad.
3. Estimaciones a partir de otras mediciones (por ejemplo, filtraciones, superficies de hoyas, etc.) empleando para ello regresiones o transformaciones de distintos tipos.

Mientras que los caudales asociados a centrales de pasada son obtenidos a partir de la función de rendimiento lineal para la central lo que, en general, no se cumple. Además, la central tiene una capacidad máxima por lo que caudales que superan esta capacidad son recortados, con lo que la serie obtenida no representará el caudal real observado. Estos datos pueden ser útiles para simular la operación del sistema, pero no son adecuados para modelos estadísticos obtenidos a partir de los datos.

2.2.3.1 Matriz de energía afluente

La matriz de energía afluente da cuenta de la energía que se podría producir en el SIC gracias a los caudales afluentes a las centrales de embalse y de pasada. Por esto los elementos de esta matriz no son los caudales afluentes, sino energías posibles de ser generadas a partir de esos caudales y que, por lo tanto, dependen de la capacidad de las centrales hidráulicas instaladas, los convenios de riego, la capacidad de aducción de las bocatomas y otras características de cada cuenca.

La energía anual afluente es la suma de las energías semanales contenidas en la matriz de energía afluente, a partir de la cual se calcula la probabilidad de excedencia de un año utilizando la expresión de Weibull como sigue:

$$Prob_{exc} = \frac{m}{n + 1}$$

Donde

m = orden del dato

n = número total de datos

En este caso se trabaja con la estadística disponible entre los años hidrológicos 1960-1961 y 2007-2008 y se le asigna el orden 1 al año con mayor energía afluente y orden 48 al año con menor energía afluente, obteniendo los resultados que se presentan en la Tabla 2-1:

Orden	Año	Energía [GWh]	Prob. de Exc.	Orden	Año	Energía [GWh]	Prob. de Exc.
1	72-73	33624.74	2.04%	25	81-82	26691.02	51.02%
2	82-83	33387.04	4.08%	26	69-70	26689.58	53.06%
3	65-66	32704.31	6.12%	27	94-95	26530.11	55.10%
4	80-81	32586.02	8.16%	28	71-72	25372.59	57.14%
5	97-98	31043.60	10.20%	29	74-75	24969.17	59.18%
6	06-07	31037.04	12.24%	30	73-74	24889.01	61.22%
7	02-03	31033.72	14.29%	31	83-84	24335.13	63.27%
8	05-06	30899.22	16.33%	32	04-05	23723.49	65.31%
9	77-78	30706.90	18.37%	33	60-61	23277.49	67.35%
10	86-87	30404.11	20.41%	34	85-86	23158.28	69.39%
11	66-67	30076.49	22.45%	35	03-04	23026.16	71.43%
12	93-94	29876.44	24.49%	36	70-71	22867.37	73.47%
13	92-93	29668.32	26.53%	37	67-68	22450.41	75.51%
14	78-79	29208.96	28.57%	38	76-77	22364.22	77.55%
15	84-85	28885.41	30.61%	39	88-89	21937.47	79.59%
16	01-02	28469.86	32.65%	40	99-00	21747.90	81.63%
17	00-01	28329.82	34.69%	41	64-65	20766.40	83.67%
18	79-80	28316.81	36.73%	42	90-91	20640.40	85.71%
19	87-88	27728.01	38.78%	43	89-90	20491.06	87.76%
20	75-76	27713.24	40.82%	44	07-08	18322.74	89.80%
21	95-96	27478.87	42.86%	45	62-63	18069.18	91.84%
22	61-62	27406.97	44.90%	46	96-97	15229.49	93.88%
23	91-92	27282.46	46.94%	47	68-69	13128.64	95.92%
24	63-64	27056.58	48.98%	48	98-99	12429.54	97.96%

TABLA 2-1: PROBABILIDAD DE EXCEDENCIA CONSIDERANDO ESTADÍSTICA DE 48 AÑOS (1960-2008)

La probabilidad de excedencia se muestra en el Gráfico 2-1:

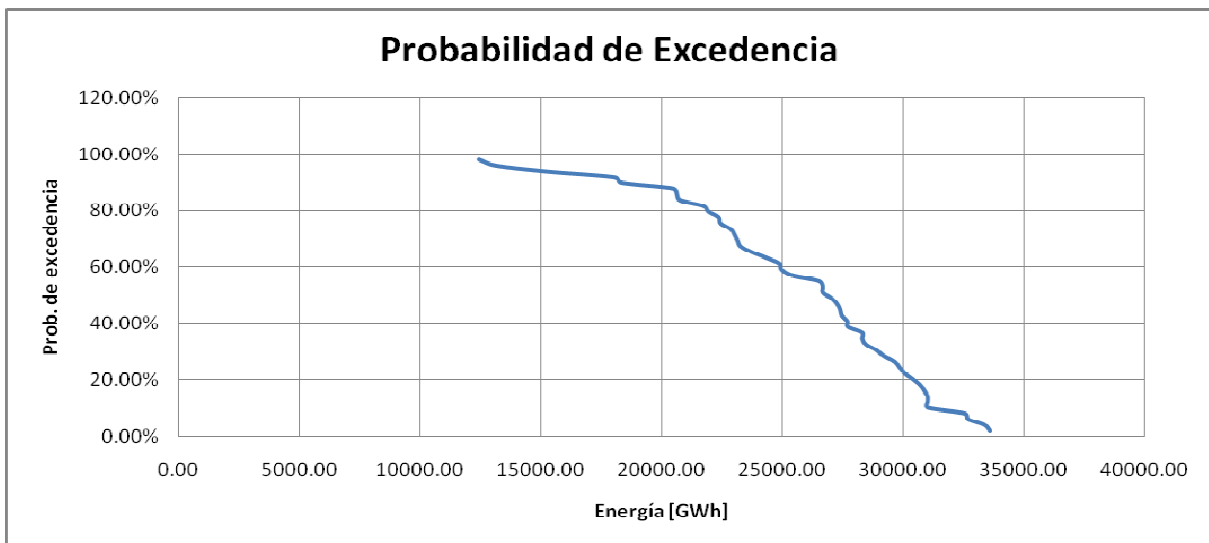


GRÁFICO 2-1: PROBABILIDAD DE EXCEDENCIA

3 SELECCIÓN DE APERTURAS

3.1 ALTERNATIVAS DE CAUDALES AFLUENTES

Entre los objetivos específicos de este trabajo de título está el proponer diversas alternativas de caudales a utilizar como aperturas de la fase de optimización de la PDDE con miras a reducir el tiempo de ejecución del PLP, pero manteniendo resultados similares a los que se obtienen actualmente utilizando 48 hidrologías históricas.

Para cumplir con este objetivo se consideran las siguientes alternativas: análisis de componentes principales, modelación de series sintéticas y agrupación de la estadística disponible en clusters.

3.1.1 ANÁLISIS DE COMPONENTES PRINCIPALES

El propósito del análisis de componentes principales (ACP o, en inglés, PCA *Principal-Component Analysis*) es reducir una muestra de datos que contienen un gran número de variables a una muestra con un número de variables reducido, pero que mantiene una parte importante de la variabilidad de la muestra original. Esto es, dadas múltiples observaciones de un número K de variables en una muestra original, se espera determinar un número M^* de esas variables que constituyan una nueva muestra. Los elementos de la nueva muestra son funciones de los datos contenidos en la original, deben contener la mayor parte de la información de ésta y M^* ser mucho menor que K . Esto es posible si existen correlaciones importantes entre las variables de la muestra original, es decir, que se tenga información redundante en la muestra. Los elementos de la nueva muestra se denominan “*componentes principales*” y constituyen una representación compacta de los datos originales, además no están correlacionados entre sí.

Usualmente, la información atmosférica y geofísica (como las series hidrológicas) exhibe una correlación importante entre variables, por lo que el ACP permite obtener una representación mucho más compacta de estas variables. Sin embargo, más allá de una simple compresión de datos el ACP puede ser una importante herramienta exploratoria en casos de gran cantidad de datos multivariable, especialmente en campos geofísicos. En estos casos, el ACP entrega importantes elementos para comprender la correlación temporal y espacial que presentan los datos a analizar.

Para mayores detalles acerca de este método de análisis estadístico se sugiere la bibliografía [12]

3.1.2 SERIES SINTÉTICAS

Las series sintéticas aparecen como una alternativa a las estadísticas históricas y han sido desarrolladas principalmente a partir del trabajo propuesto por Box y Jenkins en 1976. Box y Jenkins desarrollaron un esquema de clasificación para una gran familia de modelos de series de tiempo. En esta clasificación se distinguen los modelos Autoregresivos de orden p (AR(p)), los modelos de media móvil de orden q (MA(q)), combinaciones de los dos, llamados modelos “*autoregresivos de media móvil*” (ARMA(p,q)), y otras variantes como los modelos integrados. Sin embargo, la gran familia de modelos propuesta por Box y Jenkins, no cubre totalmente las necesidades impuestas por las series de tiempo hidrológicas, las cuales poseen algunas peculiaridades propias de los procesos geofísicos. Esto ha dado origen a una línea de investigación que ha resultado en un gran número de herramientas apropiadas para la aplicación en recursos hídricos [10].

Palacios en [10] propone un modelo de selección de aperturas para el PLP considerando que la hidrología se ha modelado con series sintéticas, de manera de asignar probabilidades de ocurrencia a los distintos escenarios hidrológicos en base a la información para el año hidrológico en curso. El supuesto del desarrollo planteado es que años hidrológicos que han presentado un comportamiento similar al año hidrológico a la fecha, deben asociarse a una mayor probabilidad de ocurrencia que aquellos que no presentan la misma variación. Con esta información es posible seleccionar aquellas condiciones hidrológicas que tengan una mayor probabilidad de ocurrencia.

3.1.3 MÉTODO DE ANÁLISIS DE CLUSTERS

El método de análisis de clusters trata de separar datos en grupos cuyas características no son conocidas en profundidad. En general, incluso el número “correcto” de clusters a obtener de una muestra es desconocido, por lo que se utiliza el grado de similitud y diferencia entre los datos para definir los grupos y asignar pertenencia a cada grupo.

Existe una gran cantidad de métodos de clustering disponibles en la actualidad. Por ejemplo, los métodos jerárquicos establecen una categoría en la agrupación de datos, de manera de reunir de a pares a aquellos grupos que se encuentran en la misma categoría, definida anteriormente. El procedimiento comienza con cada uno de los datos formando un cluster, así si hay n datos, se tendrán n grupos. El segundo paso consiste en encontrar los dos clusters (de un único elemento) que se encuentren más cercanos de acuerdo a una norma definida en el espacio de la dimensión de los datos y combinarlos. De esta manera, se tienen $(n-1)$ clusters, uno de los cuales tiene dos elementos. A continuación, en cada paso se reúnen los dos grupos más cercanos, hasta reunir los n elementos en un grupo en la iteración $(n-1)$.

Los n clusters con los que comienza el método y el único cluster que se obtiene al final son igualmente inútiles, por lo que se espera encontrar un grupo intermedio que represente los datos originales. El resultado ideal es una división de los datos que minimice la distancia entre los miembros de un mismo cluster y maximice la distancia entre miembros de diferentes clusters [12]

Como se mostró anteriormente, en un método jerárquico un dato que entra a un cluster permanece en él hasta que el algoritmo se detiene, lo que imposibilita la reubicación de puntos que pudieron haber sido mal asignados, lo cual sí es posible en los métodos no jerárquicos. Entre estos últimos, destaca el método *k-means* que utiliza un algoritmo iterativo de dos fases para minimizar la suma de la distancia entre los puntos y el centroide, sobre todos los k clusters.

La utilización de clusters como una opción para reducir el número de aperturas ha sido propuesta por Augusto en [13], donde se analiza como alternativa a la utilización de una estadística completa de 40 años hidrológicos, la agrupación de éstos en un cierto número de grupos o clusters, representados por su centroide, con una probabilidad acumulada de ocurrencia según el número de observaciones que agrupa. El análisis se basa en la comparación de los resultados de aplicar los clusters obtenidos con el método *k-means* contra los 40 años hidrológicos para resolver un problema de planificación de la operación en un modelo simplificado del SIC y en el modelo GOL, obteniéndose considerables ahorros en tiempo computacional de resolución del problema.

3.1.4 COMPARACIÓN Y SELECCIÓN

El ACP, primera opción considerada, es una buena herramienta de análisis exploratorio para una gran cantidad de datos, permitiendo obtener aquellas variables que explican la mayor parte de la varianza de la muestra. De esta manera, una muestra multivariable puede ser reducida a una con un número mucho menor de variables, lo que significará un ahorro computacional en el manejo de estos datos. Sin embargo, esta estructura no resulta compatible con el objetivo de disminuir el número de aperturas con las que trabaja el PLP, ya que el método apunta a reducir el número de variables que constituye cada dato y no el número total de datos que componen la muestra. De todas formas, no se descarta el ACP como herramienta de apoyo a las decisiones

que se puedan tomar dentro del desarrollo de la metodología de selección de aperturas a desarrollar en este trabajo.

Por otra parte las series sintéticas obtenidas a partir de modelos de series temporales, se construyen sobre la premisa que las series de tiempo tienen una historia estadística recurrente particular que puede ser modelada y explotada para fines de pronóstico. Es así como los trabajos desarrollados en torno al tema han tenido como principal objetivo el mejorar el pronóstico, es decir, obtener mejores resultados en políticas de operación al anticipar el curso de la variable hidrológica, versus al obtenido con la simulación histórica, lo cual no es el objetivo de este trabajo. Si bien Palacios en [10] propone un esquema de selección de aperturas que es compatible con el objetivo de reducir el número de éstas en la fase de optimización del PLP, se requiere que la hidrología esté modelada con series sintéticas en forma satisfactoria, lo que no fue logrado en esa oportunidad para todas las estaciones de caudales modeladas en el SIC. Es decir, para seguir el esquema propuesto debe realizarse una etapa previa de modelación de la estadística disponible, cuyos resultados no están garantizados, impidiendo que se pudiese continuar con el esquema propuesto.

Finalmente, la opción de agrupar las hidrologías históricas en clusters y disminuir el número de aperturas se ajusta al objetivo de este trabajo. Además, tiene las ventajas de que ha sido probada anteriormente con éxito en un trabajo similar y que se mantiene el formato de los datos, es decir, no se reducen variables por datos sino cantidad de datos total, lo que facilita su implementación en el esquema actual del PLP.

En vista de lo anterior, se propone reducir el número de aperturas aplicando técnicas estadísticas de clustering sobre los caudales afluentes, la metodología a seguir se muestra a continuación.

3.2 AGRUPAMIENTO DE HIDROLOGÍAS EN CLUSTERS

3.2.1 ANÁLISIS EXPLORATORIO

3.2.1.1 Metodología

El esquema de selección de aperturas que utiliza el PLP (presentado en el capítulo 2) determina qué condiciones hidrológicas se utilizarán para calcular la FCF sin diferenciar entre embalses, es decir, las aperturas definidas en los archivos de entrada del PLP son utilizadas en la fase de optimización de la PDDE de todos los embalses. Dado este carácter sistémico de la resolución del problema, se determina que se trabajará con la matriz de energía afluente del sistema, definida en 2.2.3.1, para la obtención de los clusters; pues se espera que al evaluar la energía afluente a todo el sistema, se tome la mejor decisión a nivel global.

El primer paso es determinar el número de cluster que se espera formar, en este caso, se trabaja con la siguiente progresión geométrica: 3, 6, 12, 24 y 48. La elección de este conjunto inicial está basada en la experiencia del Departamento de Planificación de la Operación del CDEC-SIC. Además, considerando que el modelo PLP debe trabajar con 48 secuencias (simulaciones/aperturas), la serie propuesta permite evaluar la progresiva reducción de las aperturas a la mitad.

Para evaluar los resultados obtenidos se proponen dos ejes principales, el primero de ellos consiste en analizar las propiedades de los clusters obtenidos, a través de la coherencia que los grupos tengan con la probabilidad de excedencia de los distintos años hidrológicos, así como las propiedades estadísticas que tiene la muestra reducida. El segundo eje es la implementación de los clusters en el esquema actual del PLP para distintos escenarios a fin de evaluar los resultados de la planificación de la operación, mediante la revisión de las siguientes variables: los costos térmicos por etapa, los costos marginales en la barra de referencia Quillota 220 kV y el valor del agua en los embalses del SIC (exceptuando Pangué y Pehuenche). La Figura 3-1 presenta esta propuesta de trabajo en forma esquemática.

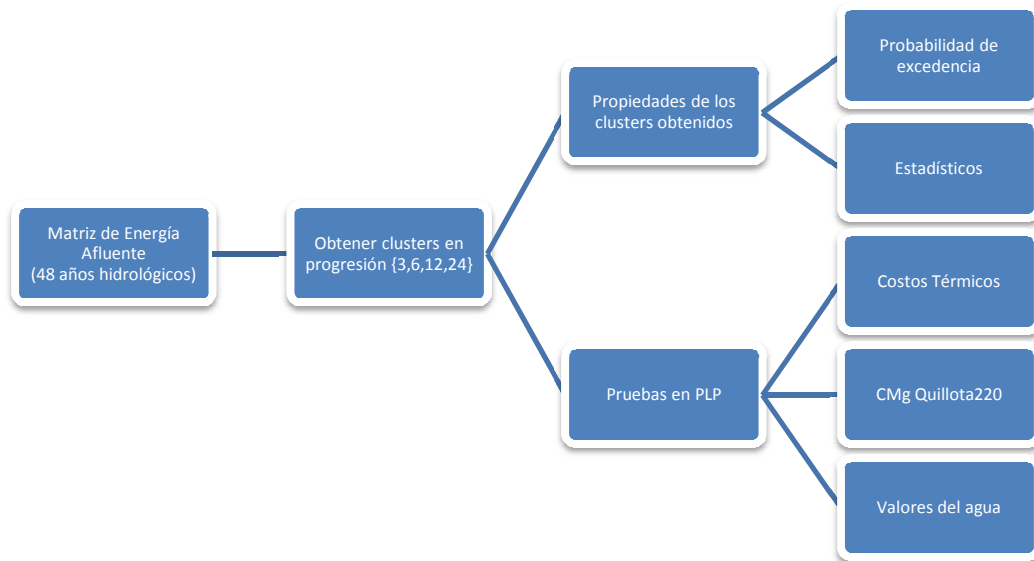


FIGURA 3-1: ESQUEMA DE TRABAJO PROPUESTO

El esquema propuesto en la Figura 3-1 constituye la primera parte de la metodología, la cual es principalmente exploratoria. De ella se espera obtener directrices para guiar el análisis y la toma de decisiones en virtud de los criterios que han sido presentados, así como de nuevos criterios que se introduzcan a lo largo del desarrollo del trabajo. De esta manera, la primera fase es una etapa general y sus resultados se evaluarán para obtener conclusiones generales; por lo que la selección del número de clusters se realiza en una segunda etapa.

3.2.1.2 Resultados y Análisis

3.2.1.2.1 Clusters Obtenidos

Como se indicó anteriormente, el primer paso a seguir es particionar la estadística disponible de 48 años hidrológicos en grupos o clusters según la progresión {3, 6, 12, 24}. Para esto se utiliza el software Matlab® de *The MathWorks™*, el cual incluye herramientas de análisis de clusters en su Statistics Toolbox™. En particular, se trabajó con el método de clustering K-means (haciendo uso de la función *kmeans*); éste es un método de clustering no jerárquico que recibe como principal parámetro el número de clusters que se desea formar con los datos².

En la Tabla 3-1 se describen otros parámetros que recibe la función *kmeans* y que son relevantes para el trabajo desarrollado.

Parámetro	Valor
<i>Distance</i>	Definición de distancia a utilizar en el espacio de los datos. La función <i>kmeans</i> minimiza con respecto a este parámetro y calcula el centroide del cluster de acuerdo a la distancia seleccionada.
<i>Replicates</i>	Número de veces que se repite el proceso de clustering, cada una a partir de un nuevo conjunto inicial (semillas). La función <i>kmeans</i> retorna la solución que minimiza la suma de distancias.
<i>Start</i>	Método utilizado para elegir las semillas (valores iniciales para los k centroides)

TABLA 3-1: PARÁMETROS DE LA FUNCIÓN *KMEANS* DE MATLAB

² Para mayor información de métodos de clustering, consultar el anexo A.

En la Tabla 3-2 se muestran los valores seleccionados para los parámetros: *distance*, *replicates* y *start*. Al seleccionar la distancia euclidiana en el espacio de 48 variables, el centroide de cada cluster es calculado como el promedio de los elementos contenidos en el cluster. De esta manera el centroide entregado por Matlab® no constituye una estadística histórica, sino un promedio de las condiciones hidrológicas que agrupa el cluster; por esta razón este resultado no puede ser utilizado directamente en la interfaz actual del PLP y se debe encontrar un valor representativo del cluster para ser utilizado en el PLP. Para esto se escoge el año hidrológico dentro del cluster que se encuentre más cercano al centroide según la distancia euclidiana; este dato es seleccionado como “*hidrología representativa*” del cluster y será el valor considerado para la apertura representativa del cluster en la implementación que se verá más adelante.

Por otra parte, se decide replicar el proceso de clustering mil veces en cada ocasión (a partir de un grupo aleatorio de semillas) para evitar que el algoritmo entregue como resultados mínimos locales en las primeras iteraciones, lo que conllevaría a errores.

Parámetro	Valor
<i>Distance</i>	Distancia euclidiana
<i>Replicates</i>	1000
<i>Start</i>	Aleatorio.

TABLA 3-2: PARÁMETROS UTILIZADOS PARA DETERMINAR LOS CLUSTERS

Se realizan las particiones propuestas del conjunto original, obteniéndose en forma exitosa cada una de ellas. A continuación se presentan los resultados de agrupar los datos en 3, 6, 12 y 24 clusters, respectivamente.

Cluster	Hidrología Representativa	Años que agrupa
1	77-78	61-62, 63-64, 65-66, 66-67, 69-70, 71-72, 72-73, 75-76, 77-78, 78-79, 79-80, 80-81, 81-82, 82-83, 84-85, 86-87, 87-88, 91-92, 92-93, 93-94, 94-95, 95-96, 97-98, 00-01, 01-02, 02-03, 05-06, 06-07
2	98-99	68-69, 96-97, 98-99
3	60-61	60-61, 62-63, 64-65, 67-68, 70-71, 73-74, 74-75, 76-77, 83-84, 85-86, 88-89, 89-90, 90-91, 99-00, 03-04, 04-05, 07-08

TABLA 3-3: RESULTADOS DE AGRUPAMIENTO EN 3 CLUSTERS

Cluster	Hidrología Representativa	Años que agrupa
1	89-90	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08
2	61-62	61-62, 63-64, 69-70, 71-72, 78-79, 79-80, 84-85, 87-88, 94-95, 95-96, 00-01
3	60-61	60-61, 70-71, 73-74, 74-75, 76-77, 83-84, 85-86, 03-04, 04-05
4	98-99	68-69, 96-97, 98-99
5	93-94	75-76, 80-81, 81-82, 86-87, 91-92, 92-93, 93-94, 01-02
6	05-06	65-66, 66-67, 72-73, 77-78, 82-83, 97-98, 02-03, 05-06, 06-07

TABLA 3-4: RESULTADOS DE AGRUPAMIENTO EN 6 CLUSTERS

Cluster	Hidrología Representativa	Años que agrupa
1	60-61	60-61, 67-68, 70-71, 74-75, 76-77, 83-84, 88-89
2	61-62	61-62, 69-70, 71-72, 78-79, 87-88, 94-95
3	84-85	63-64, 66-67, 84-85
4	06-07	95-96, 97-98, 00-01, 01-02, 06-07
5	79-80	79-80
6	93-94	86-87, 92-93, 93-94
7	98-99	68-69, 96-97, 98-99

Cluster	Hidrología Representativa	Años que agrupa
8	03-04	03-04, 04-05
9	73-74	73-74, 75-76, 85-86, 91-92
10	89-90	62-63, 64-65, 89-90, 90-91, 99-00, 07-08
11	81-82	80-81, 81-82
12	05-06	65-66, 72-73, 77-78, 82-83, 02-03, 05-06

TABLA 3-5: RESULTADOS DE AGRUPAMIENTO EN 12 CLUSTERS

Cluster	Hidrología Representativa	Años que agrupa
1	66-67	66-67, 75-76
2	72-73	72-73, 05-06
3	02-03	02-03
4	73-74	73-74, 85-86
5	06-07	65-66, 06-07
6	90-91	90-91
7	64-65	64-65, 67-68, 07-08
8	03-04	83-84, 03-04, 04-05
9	98-99	68-69, 96-97, 98-99
10	89-90	62-63, 88-89, 89-90, 99-00
11	81-82	81-82
12	79-80	79-80
13	71-72	71-72
14	93-94	86-87, 92-93, 93-94
15	61-62	61-62, 87-88
16	82-83	82-83
17	97-98	97-98
18	76-77	60-61, 70-71, 74-75, 76-77
19	78-79	77-78, 78-79, 84-85
20	63-64	63-64
21	91-92	91-92, 01-02
22	95-96	95-96, 00-01
23	80-81	80-81
24	69-70	69-70, 94-95

TABLA 3-6: RESULTADOS DE AGRUPAMIENTO EN 24 CLUSTERS

De este primer esquema de agrupamiento propuesto se desprende que los resultados son coherentes con la probabilidad de excedencia asociada a los años hidrológicos que agrupa cada cluster. Por ejemplo, en la Tabla 3-3 de agrupamiento en 3 clusters, la estadística disponible se divide en tres grupos: hidrologías húmedas, secas y medias (clusters 1, 2 y 3, respectivamente), observándose que el grupo húmedo contiene aquellos años con una probabilidad de excedencia entre 2% y 57.1%, el medio entre 59.2% y 91.85% y el seco entre 93.3% y 98%. Este resultado demuestra la conveniencia de utilizar la norma euclidiana para la formación de los clusters; ya que al utilizar otras definiciones de distancia, los resultados obtenidos discrepan de la probabilidad de excedencia de los años hidrológicos contenidos en el grupo; es decir, no es posible concluir una correlación entre la probabilidad de excedencia y los elementos que contiene el cluster.

A medida que aumenta el número de clusters en los que se divide la muestra original, las hidrologías húmedas y medias identificadas para $k = 3$ van formando clusters de menor cardinalidad. Esto da cuenta de las características de la serie a lo largo del año hidrológico, pues no sólo importa la energía afluente total anual, sino también la forma en que se va distribuyendo esa energía a lo largo del año. Es así como para el agrupamiento en 12 clusters, se tienen grupos con dos elementos e incluso el año 79-80 es el único elemento del cluster 5. Una situación similar se tiene para la partición en 24 grupos.

Un aspecto destacable es que el cluster seco (definido anteriormente para $k = 3$) se mantiene agrupado en forma intacta para todas las opciones evaluadas. Este resultado presenta la ventaja de mantener aisladas estas condiciones hidrológicas, ya que si producto del agrupamiento estas hidrológicas particularmente secas fueran reunidas con hidrológicas más húmedas se perderían parte de las condiciones extremas del sistema y disminuiría la variabilidad de la muestra con respecto a la original. Además, estos escenarios hidrológicos secos son particularmente relevantes en la planificación de operación del SIC, por lo que es importante que se mantengan en todas las opciones consideradas.

Basándose en el trabajo propuesto por Borges en [14], se decide analizar los parámetros estadísticos de las agrupaciones propuestas para compararlos con el original. Es así como se calcula para cada año representativo: la media, la desviación estándar, el coeficiente de variación y el coeficiente de asimetría de la serie, a partir de los cuales se calculan los valores promedios de la muestra conformada al agrupar las hidrológicas disponibles en clusters. Estos resultados se muestran en la Tabla 3-7, junto con los de la estadística histórica de 48 años.

k	Media [GWh]	Desviación Estándar [GWh]	Coeficiente de variación [-]	Coeficiente de asimetría [-]
3	463.9100	164.3402	0.3405	0.1864
6	504.4522	224.0021	0.4899	0.0508
12	536.6551	202.5669	0.3770	0.0513
24	561.5104	205.7031	0.3706	-0.0378
48	540.6888	201.6332	0.3772	0.0956

TABLA 3-7: ESTADÍSTICOS PARA LA MUESTRA OBTENIDA SEGÚN CADA AGRUPAMIENTO EN CLUSTERS

Puede observarse en la Tabla 3-7 que la media para todas las opciones de agrupamiento disponibles se mantiene cercana a la media 48 años hidrológicos, donde la máxima variación es de un -14.20% y se alcanza en el caso de 3 clusters. Un resultado similar se observa para la desviación estándar; sin embargo, para 3 clusters este valor se aleja en un -18.50% con respecto a la muestra original. Mientras que para el coeficiente de variación, el dato más lejano al de 48 hidrológicas es el agrupamiento en 6 clusters, lo que indica una mayor dispersión para 6 que para 48 hidrológicas. Finalmente, el coeficiente de asimetría no se mantiene al trabajar con 24, 12, 6 ó 3 hidrológicas, con lo que no es posible mantener el momento de tercer orden conforme se disminuye el tamaño de la muestra.

Una vez conseguidos los clusters, se procede a implementar las diferentes opciones de aperturas en el PLP modificando los archivos pertinentes³.

En los nuevos archivos, las aperturas seleccionadas para el período de independencia hidrológica corresponden al año representativo obtenido en el proceso de clustering; mientras que para el período de deshielo se mantuvo la optimización uni-hidrológica del esquema actual del PLP, es decir, para este período se selecciona una única apertura que coincide con el índice de la simulación en curso. Para las centrales que carecen de período de deshielo, se consignan las mismas aperturas durante todo el horizonte de planificación.

³ PLPIDAPE.DAT y PLPIDAP2.DAT

3.2.1.2.2 Resultados PLP

La Dirección de Operación del CDEC-SIC realiza periódicamente (al menos una vez a la semana) la planificación de mediano – largo plazo, denominada Programación Semanal, la cual tiene por objetivo determinar listas de mérito de las diversas centrales del Sistema Interconectado Central con las cuales se realizará el despacho en tiempo real. Para esto, se ejecuta el modelo PLP en un horizonte aproximado de dos años, con etapas semanales para las primeras cuatro semanas y mensuales desde la quinta semana hasta el final del horizonte. En las etapas semanales, se simula una curva de duración del consumo mediante 5 bloques, mientras que para el caso de las etapas mensuales, se utiliza sólo un bloque de consumo. El sistema de transmisión se modela hasta niveles en torno a los 110 kV y en algunas zonas hasta 66 kV.

El esquema de modificación de aperturas se implementó en dos programaciones semanales: la que comienza el 17 de abril de 2009 y la que se inicia el 28 de agosto de 2009. Los resultados del primer caso se presentan a continuación, mientras que los del segundo se incluyen en el anexo. En todos los casos se trabajó con 48 hidrologías en la etapa de simulación, es decir, se consideran todas las hidrologías históricas. Esto resulta relevante, considerando que el reglamento del CDEC-SIC estipula que en los modelos de Planificación de la Operación se considere la estadística hidrológica utilizada por la Comisión Nacional de Energía en la fijación de precio de nudo que se encuentre vigente.

La correspondencia entre las etapas y las fechas del horizonte de planificación, se muestran en el anexo B.

3.2.1.2.2.1 Tiempos de Ejecución

La ejecución del PLP se lleva a cabo en un entorno Linux, utilizando un computador Intel® Pentium® D CPU 3.40 GHz con RAM de 2.0 GB. El primer resultado que se analiza corresponde a los tiempos de ejecución alcanzados para cada una de las pruebas y que se presentan a continuación:

Número de aperturas	Número de simulaciones	Tiempo total
3	48	00:11:34
6	48	00:39:58
12	48	00:53:42
24	48	01:20:35
48	48	02:09:15

TABLA 3-8: TIEMPOS DE EJECUCIÓN SEGUN NÚMERO DE APERTURAS

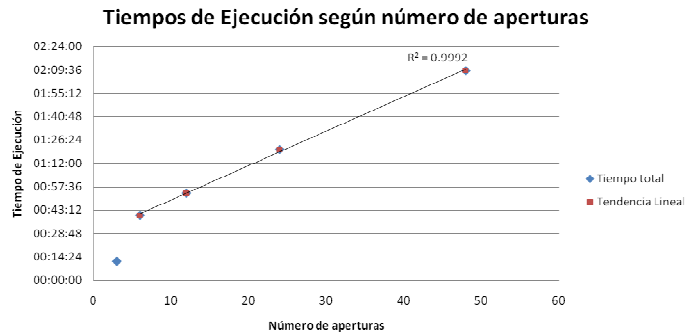


GRÁFICO 3-1: TIEMPO DE EJECUCIÓN PLP SEGÚN NÚMERO DE APERTURAS

A partir de la Tabla 3-8 y del Gráfico 3-1, puede observarse una relación lineal entre el número de aperturas y el tiempo total de ejecución del PLP; evidenciándose una reducción sistemática desde el caso base que demora alrededor de dos horas para 48 aperturas, a casi 40 minutos para 6 aperturas. Tanto en el caso de Abril, como en Agosto, si se utilizan 3 aperturas el programa no finaliza alcanzando el número máximo de iteraciones, sino por ajustarse erróneamente al umbral de convergencia predefinido, en 8 y 5 iteraciones, respectivamente; por esto el tiempo de ejecución en estos casos se escapa de la tendencia lineal expresada anteriormente. Como información adicional se agrega que convergencias obtenidas en muy pocas iteraciones en el PLP para el problema de coordinación hidrotérmica en estudio, adolece de inconsistencias en sus resultados puntuales en varias etapas.

3.2.1.2.2 Valores Esperados

Para evaluar la calidad de los resultados obtenidos al reducir el número de aperturas en la etapa de optimización del PLP se decide utilizar las series de los costos térmicos, los costos marginales en Quillota 220 kV y el valor del agua en los embalses del SIC (exceptuando Pangué y Pehuenche), dada la importancia de estas variables en los procesos de Planificación de la Operación del CDEC-SIC. Las series obtenidas se contrastaron con el caso base definido como el escenario con 48 aperturas y 48 simulaciones.

En el Gráfico 3-2 se muestran las series obtenidas para el valor esperado de los costos marginales en la barra Quillota 220 kV, es decir, para el promedio de las simulaciones obtenidas en el caso del 17/04/2009. En él es posible observar que los resultados de 24, 12 y 6 aperturas se ajustan al caso base sin dispersión significativa, esto es, que en cada etapa el valor obtenido resulta cercano al valor base de 48 aperturas. La mayor variación se tiene en la etapa 32 correspondiente a abril del 2010, en donde el valor obtenido con 6 aperturas resulta un 8.24% superior al caso base. En cambio, el resultado obtenido usando 3 aperturas difiere al caso base en reiteradas ocasiones dentro del horizonte de planificación, llegando a ser un 21.84% superior en la primera etapa, lo que imposibilita cualquier aplicación de estos resultados dentro de los procesos de Planificación de la Operación; éste es un resultado esperado considerando que el programa no terminó por número de iteraciones sino por ajustarse erróneamente al umbral de convergencia en 8 iteraciones. El lector puede consultar el anexo para revisar los resultados del valor esperado para el caso de 28/08/2009, en el cual se obtienen apreciaciones similares.

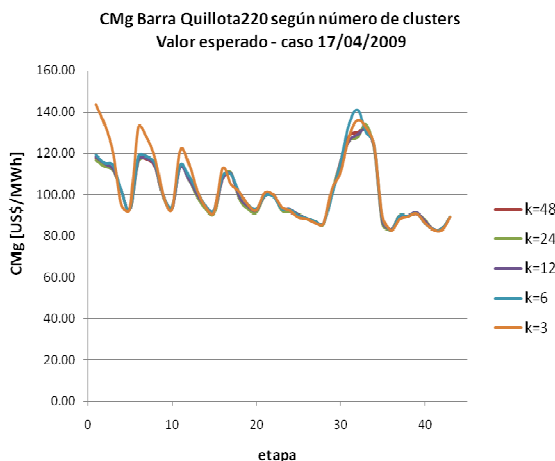


GRÁFICO 3-2: COSTOS MARGINALES EN LA BARRA QUILLOTA 220, VALOR ESPERADO PARA EL CASO 17/04/2008

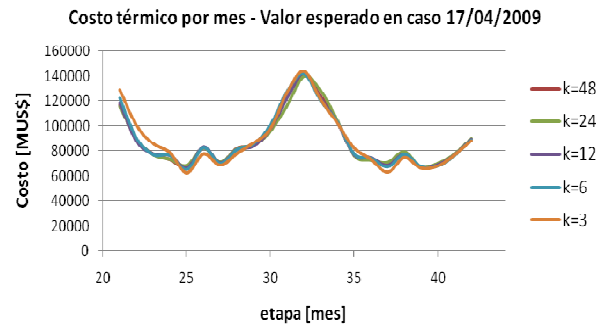


GRÁFICO 3-3: COSTOS TÉRMICOS POR ETAPAS, VALOR ESPERADO PARA EL CASO DEL 17/04/2009

Otro aspecto a destacar es que al reducir el número de aperturas no se introduce un sesgo sistemático sobre los resultados, pues la variación que se obtiene con respecto a utilizar 48 aperturas dependerá del caso y del número de aperturas.

A diferencia de los costos marginales que son obtenidos directamente como un resultado del PLP, los costos térmicos por etapa resultan de multiplicar la energía generada de cada central térmica dentro de una etapa, por el costo variable de la central en esa etapa. Debido a que las primeras etapas del horizonte de planificación son de duración variable, los gráficos se presentan agrupando las primeras etapas a fin de mostrar el costo térmico mensual. Estos valores se encuentran actualizados a una tasa del 10% anual.

El resultado es similar al obtenido para los costos marginales, ya que el valor esperado de los costos térmicos por etapa, tanto para el caso de Abril como para el de Agosto, resultan cercanos al caso base de 48 aperturas utilizando 6, 12 ó 24 aperturas. De igual manera, el caso de 3 aperturas es el que presenta una mayor

dispersión, por ejemplo, en el caso del Gráfico 3-2, la máxima diferencia se produce en la etapa 27 y equivale a un +10.95%. Por otra parte, al igual que para los costos marginales en la barra Quillota 220 kV, se tiene que al reducir el número de aperturas no se introduce un sesgo sistemático sobre los resultados, pues la variación que se obtiene con respecto a utilizar 48 aperturas va cambiando etapa a etapa como puede observarse en los gráficos anteriores.

Finalmente, se evalúan los valores del agua de los embalses del SIC. Al igual que los costos marginales en la barra Quillota 220 kV, los valores del agua en los embalses son un resultado directo del PLP. Los resultados de los embalses Pangue y Pehuenche no se consideran para la evaluación, debido a que son embalses de menor capacidad que pueden alcanzar más fácilmente su cota mínima o máxima dentro del horizonte de planificación y, al hacerlo, desacoplan el problema de coordinación hidrotérmica en el tiempo, por lo que se pierde el interés en este tipo de resultados. De esta manera, los embalses a considerar son: Canutillar, Cipreses (Laguna La Invernada), Colbún, el Toro (Lago Laja), Laguna del Maule, Ralco y Rapel.

El valor del agua obtenido al final de la primera semana del horizonte de planificación es particularmente relevante dentro de los procesos de Planificación de la Operación del CDEC-SIC y es una condición de borde que recibe la Programación de Corto Plazo (PCP) a partir del PLP; siendo el resultado que tiene una aplicación más directa, en comparación con los costos marginales o los costos térmicos. A continuación se presentan los resultados del valor del agua para la primera semana en el caso de abril, así como su variación porcentual con respecto a utilizar 48 aperturas.

Embalse	Valor del agua al final de la primera semana [US\$/dam ³] y variación porcentual				
	48	24	12	6	3
Canutillar	63.59	63.5 (-0.14%)	64.87 (2.01%)	65.38 (2.81%)	69.26 (8.92%)
Cipreses	266.10	263.69 (-0.91%)	264.46 (-0.62%)	270.07 (1.49%)	287.47 (8.03%)
Colbún	63.06	62.41 (-1.03%)	62.45 (-0.97%)	63.7 (1.01%)	70.69 (12.1%)
El Toro	126.52	119.9 (-5.23%)	131.63 (4.04%)	134.33 (6.17%)	155.79 (23.13%)
Laguna del Maule	75.43	75.43 (0%)	75.43 (0%)	75.42 (-0.01%)	75.41 (-0.03%)
Ralco	73.45	72.87 (-0.79%)	73.23 (-0.3%)	74.47 (1.39%)	82.77 (12.69%)
Rapel	20.17	19.89 (-1.39%)	20.35 (0.89%)	20.75 (2.88%)	23.06 (14.33%)

TABLA 3-9: VALOR DEL AGUA AL FINAL DE LA PRIMERA SEMANA [US\$/DAM³] Y VARIACIÓN PORCENTUAL PARA EL CASO DE 17/04/2009

Estos resultados muestran que el valor del agua en algunos embalses es sensible al número de aperturas utilizado en la etapa de optimización del PLP, particularmente Canutillar, Cipreses y Colbún. Sin embargo, no se establece un sesgo sistemático, teniéndose variaciones positivas y negativas. Nuevamente, se comprueba que los resultados obtenidos utilizando 3 aperturas son significativamente distintos a los obtenidos con 48 aperturas; mientras que utilizando 6, 12 ó 24 aperturas los resultados obtenidos no difieren en más de un $\pm 10\%$ con respecto al caso base.

Junto con analizar el valor del agua al final de la primera semana del horizonte de planificación, es relevante revisar cómo se comportan las series obtenidas a lo largo de este horizonte. Es por esto que se presenta la siguiente desagregación por embalse para 6, 12 y 24 aperturas. La opción de 3 aperturas se descarta a partir de este punto, pues ya se ha constatado que no es una alternativa viable de acuerdo a lo señalado anteriormente.

Canutillar

Los resultados obtenidos para el valor esperado del caso de abril muestran que para 24 y 12 aperturas, las series se asemejan al caso base y no se observa diferencias superiores al $\pm 5\%$, mientras que para 6 aperturas, se tienen etapas en las cuales el valor obtenido es entre un 5% y 6% mayor al caso base. Para el caso de agosto, los resultados de 24 aperturas son coincidentes con el caso base, mientras que los obtenidos para 12 y 6 aperturas muestran una dispersión de +7.6% y +9.7% respectivamente.

Cipreses

El valor del agua esperado en el embalse Cipreses para los casos de abril y de agosto, se ajusta al caso base en las primeras 20 etapas del horizonte de planificación ya sea utilizando 24, 12 ó 6 aperturas. Sin embargo a partir de las etapas mensuales, los resultados pueden mostrar etapas puntuales donde el valor obtenido con un menor número de aperturas es notoriamente distinto al caso base.

Colbún

Para el valor esperado en el caso de abril las series obtenidas para 24 y 12 aperturas se asemejan al caso base salvo etapas puntuales, mientras que para 6 aperturas se presenta un alejamiento promedio de +11.3% entre las etapas 11 y 21, lo que corresponde desde la tercera semana de abril hasta mayo de 2009. Para el caso de agosto, los resultados obtenidos son mejores, pues para 24 y 12 aperturas no hay variaciones importantes con respecto al caso base; mientras que para 6 aperturas, la dispersión promedio con respecto al caso base en las primeras etapas es de +8.2%.

El Toro

Para el caso de abril, las series obtenidas para el valor esperado difieren del caso base en promedio en un -5.3% para 24 aperturas y en un +5.9% para 6 aperturas, mientras que para 12 aperturas se observa un valor similar al del caso base. Por otra parte, el caso de agosto no muestra diferencias relevantes entre utilizar 48, 24, 12 ó 6 aperturas.

Laguna del Maule

Para la Laguna del Maule se tiene que, independiente del mes en que simule o del número de aperturas considerado, los resultados obtenidos son similares a los de 48 aperturas y no hay dispersión superior al $\pm 5\%$.

Ralco

El valor esperado de la serie se ajusta al caso base en el caso de abril para 24, 12 y 6 aperturas; para agosto, se tienen resultados similares al de 48 aperturas en 24 y 12 aperturas, pero en 6 aperturas se tiene una dispersión promedio de +5% con respecto al caso base.

Rapel

Los valores esperados para abril y agosto son satisfactorios para 24, 12 y 6 aperturas, ya que no se observan diferencias mayores a un $\pm 5\%$ con respecto a utilizar 48 aperturas.

Al finalizar esta primera etapa se puede concluir que la reducción de aperturas involucra una disminución del tiempo de ejecución del PLP en forma lineal. Por otro lado, el número mínimo de aperturas a utilizar es 6, ya que el modelo entrega una solución incorrecta para 3 aperturas, mientras que para 6, 12 y 24 se obtienen resultados en el modelo PLP que son similares a los obtenidos al utilizar 48 aperturas. Dado que al utilizar entre 6 y 12 aperturas existe una importante posibilidad de reducción en los tiempos de ejecución, manteniendo la calidad de los resultados, se trata de encontrar parámetros más finos para definir el número final de aperturas a utilizar, ya sea basándose en las características estadística de la muestra o en los resultados entregados por el modelo PLP.

3.2.1.2.2.3 Simulaciones de años particulares

Con el fin de observar qué ocurre con los resultados de los casos anteriores para simulaciones que representan años hidrológicos específicos, se seleccionan 3 simulaciones representativas de condiciones hidrológicas: húmeda, media y seca. Para el caso de abril se considera la Matriz de Energía Afluente definida en 2.2.3.1 a fin de evaluar la energía afluente en todo el año hidrológico y así escoger las simulaciones que representen las distintas condiciones hidrológicas señaladas, pues habrá un período considerable del año hidrológico seleccionado dentro del horizonte de planificación. Con ese antecedente se determinan los años con probabilidad de excedencia 20%, 50% y 90% como años: húmedo, medio y seco, respectivamente, los cuales se muestran en la tabla a continuación:

Año Hidrológico	Probabilidad de excedencia	Condición
86-87	20.41%	Húmedo
81-82	51.02%	Medio
07-08	89.80%	Seco

TABLA 3-10: CONDICIONES HIDROLÓGICAS SELECCIONADAS DE ACUERDO A SU PROBABILIDAD DE EXCEDENCIA

Una vez que han sido definidos los escenarios hidrológicos considerados como: húmedo, medio y seco, se presentan las series de costos marginales en Quillota 220 kV para 6, 12, 24 y 48 aperturas.

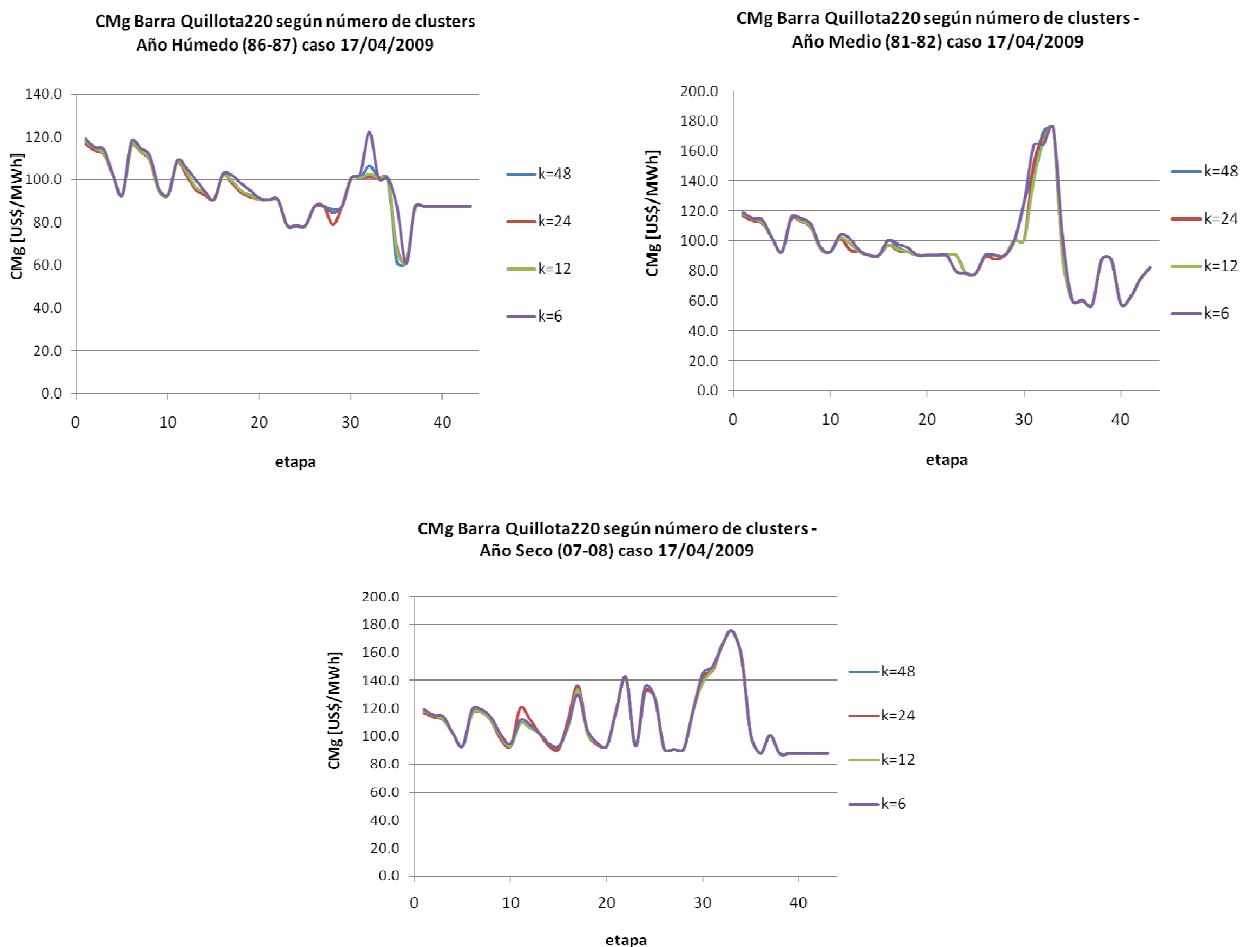


FIGURA 3-2: GRÁFICOS DE COSTOS MARGINALES PARA AÑOS: HÚMEDO, MEDIO Y SECO

Para el año húmedo puede verse, que al igual que para el promedio de las simulaciones, los costos marginales para 24, 12 y 6 aperturas se asemejan a la serie de 48 aperturas. No se observan dispersiones importantes en las primeras etapas del horizonte de planificación, pero es relevante notar que para 6 aperturas se produce una diferencia puntual en la etapa de julio de 2010 (etapa 35), que equivale +42.81%; sin embargo, se produce en una etapa mensual cercana al final del horizonte de planificación.

Una situación similar se tiene para el año medio, pues las alternativas evaluadas entregan resultados similares al caso base durante todo el horizonte de planificación, observándose sólo diferencias puntuales en las etapas mensuales. Estas diferencias alcanzan un valor máximo de -20.12% para 24 y 12 aperturas en la etapa 30, que también corresponde a una etapa mensual.

Con respecto al año seco, las tres alternativas responden adecuadamente con respecto al caso base, teniéndose un panorama análogo al de los años: húmedo y medio. Este es un resultado muy favorable, pues las condiciones más extremas, en este caso, los años hidrológicos más secos son relevantes para la planificación de la operación. Por ende, el que la simulación no muestre diferencias mayores al $\pm 5\%$ para 24, 12 ó 6 aperturas es particularmente positivo.

Finalmente, se destaca que en las simulaciones seleccionadas no se produce un sesgo sistemático por la reducción de aperturas

3.2.2 EVALUACIÓN DE ALTERNATIVAS DE AGRUPAMIENTO

3.2.2.1 Metodología

A partir de los resultados abordados se amplía el método de trabajo propuesto inicialmente en el punto 3.1.1, ya que se propone refinar la búsqueda del número de aperturas en torno a 6 y 12. Para esto se plantea un nuevo esquema de trabajo a fin de explorar aspectos que no habían sido tratados o bien profundizar el esquema anterior.

El primer punto a considerar tiene que ver con la cardinalidad de los clusters y probabilidad asociada. Actualmente el modelo PLP trabaja los escenarios hidrológicos como eventos equiprobables, asignando a cada una de las 48 hidrologías históricas la misma probabilidad de ocurrencia. Al calcular la función de costo futuro, los resultados para cada hidrología (en la etapa de independencia hidrológica) son promediados para obtener el valor de cada estado y etapa. Por lo anterior, al reducir el número de aperturas se rompe esta equiprobabilidad asignando una mayor probabilidad de ocurrencia a las hidrologías representativas de cada clúster. En particular, como se describió en 3.2.1.2.1 existe un cluster que agrupa a las tres hidrologías más secas y que se mantiene para todas las opciones de partición evaluadas. Así, al disminuir el número de aperturas, se produce la agrupación de las hidrologías consideradas medias y húmedas, mientras que el cluster seco se mantiene particionado del resto aumentando su probabilidad de ocurrencia desde un 6.3% (3 aperturas de 48) hasta 16.7% (una apertura de 6). A fin de evitar este fenómeno, se evalúa la cardinalidad de los cluster, para encontrar la mínima que se pueda asociar a un número de aperturas equivalente a la probabilidad acumulada. Es decir, si un cluster agrupa 3 hidrologías y otro cluster agrupa 9 hidrologías, es natural pensar que la representación de la muestra con respecto a la original debería ponderar al segundo por 3 en comparación con el primero. Para eso se propone evaluar la cardinalidad de los clusters entre 6 y 12, a fin de encontrar una partición que permita emular escenarios equiprobables en base a la cardinalidad de los grupos encontrados.

Como ya se vio que entre 12 y 6 existe una importante posibilidad de reducción en los tiempos de ejecución, manteniendo la calidad de los resultados, en este punto se trata de encontrar parámetros más finos para definir el número final de aperturas a utilizar, ya sea basándose en las características estadística de la muestra o en los resultados del PLP, cuando se utilizan entre 6 y 12 aperturas.

3.2.2.2 Resultados y Análisis

3.2.2.2.1 Clusters Obtenidos

Para particionar la estadística histórica se procede de la misma forma que en 3.2.1, a partir de lo cual se obtienen los siguientes resultados:

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
1	93-94	5	80-81, 81-82, 86-87, 92-93, 93-94
2	61-62	5	61-62, 63-64, 79-80, 84-85, 02-03
3	94-95	10	69-70, 71-72, 75-76, 78-79, 87-88, 91-92, 94-95, 95-96, 00-01, 01-02
4	05-06	8	65-66, 66-67, 72-73, 77-78, 82-83, 97-98, 05-06, 06-07
5	60-61	9	60-61, 70-71, 73-74, 74-75, 76-77, 83-84, 85-86, 03-04, 04-05
6	98-99	3	68-69, 96-97, 98-99
7	89-90	8	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08

TABLA 3-11: RESULTADOS DE AGRUPAMIENTO EN 7 CLUSTERS

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
1	98-99	3	68-69, 96-97, 98-99
2	02-03	4	63-64, 79-80, 84-85, 02-03
3	60-61	7	60-61, 70-71, 74-75, 76-77, 83-84, 03-04, 04-05
4	05-06	8	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08
5	93-94	5	80-81, 81-82, 86-87, 92-93, 93-94
6	94-95	9	61-62, 69-70, 71-72, 78-79, 87-88, 94-95, 95-96, 00-01, 01-02
7	05-06	8	65-66, 66-67, 72-73, 77-78, 82-83, 97-98, 05-06, 06-07
8	73-74	4	73-74, 75-76, 85-86, 91-92

TABLA 3-12: RESULTADOS DE AGRUPAMIENTO EN 8 CLUSTERS

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
1	79-80	2	79-80, 02-03
2	84-85	3	63-64, 66-67, 84-85
3	05-06	7	65-66, 72-73, 77-78, 82-83, 97-98, 05-06, 06-07
4	94-95	9	61-62, 69-70, 71-72, 78-79, 87-88, 94-95, 95-96, 00-01, 01-02
5	93-94	5	80-81, 81-82, 86-87, 92-93, 93-94
6	89-90	8	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08
7	60-61	7	60-61, 70-71, 74-75, 76-77, 83-84, 03-04, 04-05
8	98-99	3	68-69, 96-97, 98-99
9	73-74	4	73-74, 75-76, 85-86, 91-92

TABLA 3-13: RESULTADOS DE AGRUPAMIENTO EN 9 CLUSTERS

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
1	84-85	3	63-64, 66-67, 84-85
2	06-07	5	95-96, 97-98, 00-01, 01-02, 06-07
3	98-99	3	68-69, 96-97, 98-99
4	89-90	8	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08
5	93-94	5	80-81, 81-82, 86-87, 92-93, 93-94

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
6	05-06	6	65-66, 72-73, 77-78, 82-83, 02-03, 05-06
7	61-62	6	61-62, 69-70, 71-72, 78-79, 87-88, 94-95
8	79-80	1	79-80
9	60-61	7	60-61, 70-71, 74-75, 76-77, 83-84, 03-04, 04-05
10	73-74	4	73-74, 75-76, 85-86, 91-92

TABLA 3-14: RESULTADOS DE AGRUPAMIENTO EN 10 CLUSTERS

N° Cluster	Año Hidrológico	Cardinalidad	Años hidrológicos que agrupa
1	84-85	3	63-64, 66-67, 84-85
2	79-80	1	79-80
3	89-90	8	62-63, 64-65, 67-68, 88-89, 89-90, 90-91, 99-00, 07-08
4	93-94	3	86-87, 92-93, 93-94
5	06-07	5	95-96, 97-98, 00-01, 01-02, 06-07
6	81-82	2	80-81, 81-82
7	05-06	6	65-66, 72-73, 77-78, 82-83, 02-03, 05-06
8	61-62	6	61-62, 69-70, 71-72, 78-79, 87-88, 94-95
9	73-74	4	73-74, 75-76, 85-86, 91-92
10	60-61	7	60-61, 70-71, 74-75, 76-77, 83-84, 03-04, 04-05
11	98-99	3	68-69, 96-97, 98-99

TABLA 3-15: RESULTADOS DE AGRUPAMIENTO EN 11 CLUSTERS

Como se indicó anteriormente, a partir de estos clusters se espera relacionar la cardinalidad y la probabilidad, por lo que en la Tabla 3-16, se presenta en forma resumida la cardinalidad mínima y máxima que se observan en las distintas opciones de agrupamiento propuestas.

k	Cardinalidad de grupos formados	
	Mínima	Máxima
6	3	11
7	3	10
8	3	9
9	2	9
10	1	8
11	1	8
12	1	7

TABLA 3-16: CARDINALIDAD MÍNIMA Y MÁXIMA AL PARTICIONAR ENTRE 6 Y 12.

De la tabla se desprende que a medida que aumenta el número de clusters, la cardinalidad mínima y máxima de los grupos disminuye, dando cuenta del aumento en la dispersión del agrupamiento de la muestra. También se observa de la Tabla 3-11 a Tabla 3-15, que el cluster que agrupa las hidrologías más secas se mantiene sin alteraciones en todas las alternativas propuestas, lo que refuerza la idea de trabajar con una cardinalidad mínima de 3.

Por otra parte, en la Tabla 3-17 se muestra la media de la muestra calculada en forma simple y en forma ponderada (de acuerdo al número de elementos que incluye cada cluster). Se espera que al obtener la media ponderada se corrija el efecto que tiene el cluster seco sobre la selección de aperturas, acercando el promedio de la selección al promedio de las 48 hidrologías históricas, lo cual se tiene para 6, 7, 9 y 12 aperturas; mientras que para 8, 10 y 11 se tiene el efecto contrario, es decir, la media no ponderada se acerca más al promedio de las 48 hidrologías históricas.

k	Promedio Simple [GWh]	Promedio Ponderado [GWh]
6	504.4522	536.9091
7	511.8552	528.3180
8	549.9950	570.7197
9	525.4463	529.1215
10	539.8288	538.5984
11	541.6112	535.8202
12	536.6551	538.0009
48	540.6888	540.6888

TABLA 3-17: MEDIAS SIMPLE Y PONDERADA SEGÚN NÚMERO DE CLUSTERS

A partir de estos resultados se decide evaluar dos nuevas alternativas: definir nuevas aperturas para incluir la noción de equiprobabilidad y seleccionar el número de aperturas de acuerdo a la media de la muestra seleccionada.

Para la primera opción se trabajará con 6 clusters, a partir de esta muestra se definen 17 aperturas, considerando la cardinalidad mínima de 3, es decir, a cada cluster se le asigna una apertura cada 3 elementos contenidos en él, en caso que la cardinalidad del cluster no sea divisible por 3 se aproxima al entero más próximo para efectos de definir el número de repeticiones. No obstante, se han determinado un mayor número de aperturas por cluster, todas las aperturas que se asignan a cada cluster corresponderán al año hidrológico representativo definido de acuerdo al criterio establecido en el punto 3.2.1.2.1 (cercanía al centroide). Se toma esta decisión, considerando que estas opciones continúan siendo exploratorias y este caso se puede contrastar con la opción ya evaluada de 6 clusters original con sólo 6 aperturas.

La segunda opción a evaluar será escoger el número de aperturas que determine una selección con valor promedio lo más cercano a la estadística histórica. De esta manera, se seleccionan 10 aperturas cuyo valor promedio difiere en un -0.16% con respecto a los 48 años hidrológicos.

3.2.2.2 Resultados PLP

3.2.2.2.1 Tiempos de ejecución

Con respecto a los tiempos de ejecución se tiene que los casos de 6, 10, 12, 24 y 48 aperturas se ajustan a la tendencia lineal mostrada anteriormente, obteniéndose un coeficiente de correlación de 0.9993, como puede apreciarse en el Gráfico 3-4. Sin embargo, el caso de 17 aperturas tiene un tiempo de ejecución que escapa de la tendencia lineal y es comparable con el tiempo correspondiente a 6 aperturas. Cabe recordar que las 17 aperturas fueron obtenidas a partir de la partición en 6 grupos, pero replicando el año hidrológico representativo de acuerdo a la cardinalidad de cada cluster.

Nº de aperturas	Nº de simulaciones	Tiempo total
6	48	00:39:58
10	48	00:49:01
12	48	00:53:42
17	48	00:43:20
24	48	01:20:35
48	48	02:09:15

TABLA 3-18: TIEMPOS DE EJECUCIÓN

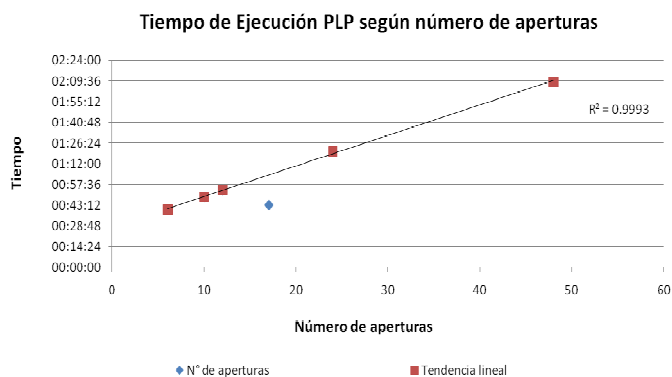


GRÁFICO 3-4: TIEMPO DE EJECUCIÓN PLP SEGÚN NÚMERO DE APERTURAS

3.2.2.2.2 Valores esperados

Los valores esperados de costos marginales en la barra Quillota 220 kV y de costos térmicos por etapa no presentan diferencias relevantes del caso base para todas las alternativas. Por una parte para el costo marginal esperado se evidencian diferencias en de +8.24% y +6.43% en la etapa mensual de abril de 2010 (etapa 32) para 6 y 17 aperturas; mientras que el costo térmico esperado no muestra diferencias superiores al $\pm 5\%$ con respecto al caso de 48 aperturas.

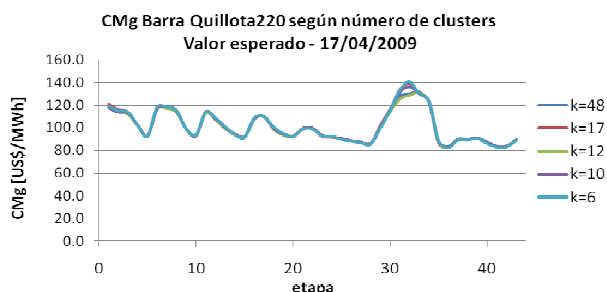


GRÁFICO 3-5: COSTOS MARGINALES EN LA BARRA QUILLOTA 220, VALOR ESPERADO PARA EL CASO DEL 17/04/2009

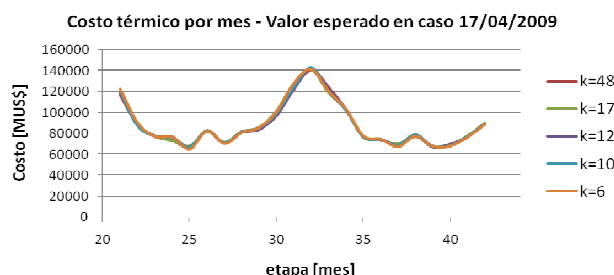


GRÁFICO 3-6: COSTOS TÉRMICOS POR ETAPAS, VALOR ESPERADO PARA EL CASO DEL 17/04/2009

De esta manera, al analizar estas nuevas alternativas: replicación de aperturas para introducir nociones de equiprobabilidad versus determinar la muestra que tiene una media similar a la hidrología histórica, se puede ver que ambas opciones son comparables en la calidad de sus resultados y no hay evidencia que permita preferir una opción sobre otra. Situación que se repite en las simulaciones específicas de años particulares.

El valor de agua en los embalses refuerza la idea de que todas las opciones contempladas son válidas y ambas mejoran el resultado de 6 aperturas, acercándose más al valor de 48 aperturas. Si bien es el caso de 10 aperturas el que presenta menor dispersión en todos los embalses al analizar los datos contenidos en la Tabla 3-19, esto dependerá del caso particular en que se esté evaluando la ejecución del modelo PLP (por ejemplo, Abril o Agosto), por lo que no hay evidencia significativa que permita seleccionar una alternativa sobre la otra.

Embalse	Valor del agua al final de la primera semana [US\$/dam ³] y variación porcentual para caso 17/04/2009				
	48	17	12	10	6
Canutillar	63.59	64.72 (1.78%)	64.87 (2.01%)	64.95 (2.14%)	65.38 (2.81%)
Cipreses	266.10	269.59 (1.31%)	264.46 (-0.62%)	265.7 (-0.15%)	270.07 (1.49%)
Colbún	63.06	64.08 (1.62%)	62.45 (-0.97%)	62.95 (-0.17%)	63.7 (1.01%)
El Toro	126.52	117.81 (-6.88%)	131.63 (4.04%)	136.21 (7.66%)	134.33 (6.17%)
Laguna del Maule	75.43	75.43 (0%)	75.43 (0%)	75.43 (0%)	75.42 (-0.01%)
Ralco	73.45	74.25 (1.09%)	73.23 (-0.3%)	73.64 (0.26%)	74.47 (1.39%)
Rapel	20.17	20.22 (0.25%)	20.35 (0.89%)	19.97 (-0.99%)	20.75 (2.88%)

TABLA 3-19: VALOR DEL AGUA AL FINAL DE LA PRIMERA SEMANA [US\$/dam³] Y VARIACIÓN PORCENTUAL PARA EL CASO DE 17/04/2009

Por lo tanto, esta segunda evaluación de alternativas entrega como conclusión relevante que es posible acotar el número de aperturas en un rango entre 6 y 12; sin embargo, el análisis no determina la preferencia de una opción sobre otra, ya que la calidad de los resultados obtenidos se encuentra dentro de los rangos considerados como satisfactorios para todas las opciones, tanto para el caso de Abril, como para el de Agosto. Es por esto que a continuación se presentan criterios de selección del número de aperturas a utilizar coherentes con la metodología que se ha desarrollado de agrupación de hidrologías en clusters.

3.2.3 CRITERIOS DE SELECCIÓN

3.2.3.1 Tiempos de Ejecución

El primero de los criterios de selección propuesto tiene que ver con los tiempos de ejecución del modelo PLP, ya que dentro de los objetivos principales del trabajo de memoria se encuentra el encontrar alternativas que permitan la reducción de estos tiempos, pero manteniendo la precisión de los resultados con respecto a la situación actual. Es por eso que se ha evaluado el tiempo de ejecución de acuerdo al número de aperturas en los puntos anteriores, observándose en primera instancia una relación lineal entre el número de aperturas y el tiempo; pero en 3.2.2 se presenta un caso que rompe con esa tendencia. Debido a esto, se desea explorar la posibilidad de aumentar el número de aperturas sin aumentar el tiempo de ejecución, en la medida que se replican las hidrologías utilizadas. Es así como se realizan más simulaciones de casos similares cuyos resultados se muestran en Tabla 3-20.

Nº de aperturas	Nº de Clusters	Tiempo total
6	1	00:29:10
6	6	00:39:58
10	10	00:49:01
12	12	00:53:42
12	6	00:39:50
17	6	00:43:20
24	24	01:20:35
48	48	02:09:15

TABLA 3-20: TIEMPOS DE EJECUCIÓN

En la Tabla 3-20 se incluyen el número de aperturas y el número de clusters a partir del cual se obtiene la muestra. En los casos en los que el número de aperturas y el número de clusters coinciden, significa que cada cluster fue representado mediante una apertura (hidrología representativa); mientras que cuando estos números difieren, se asigna al menos una apertura por cluster, existiendo clusters con más aperturas. A partir de esta tabla se generan los gráficos 3-7 y 3-8.

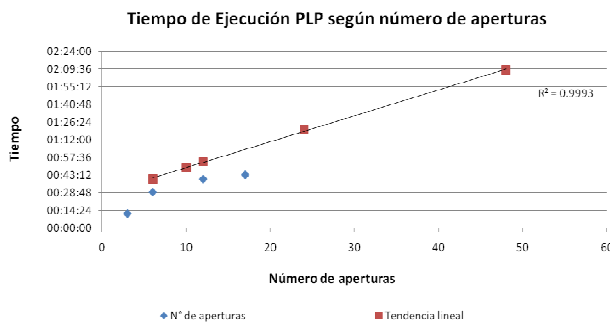


GRÁFICO 3-7: TIEMPOS DE EJECUCIÓN SEGÚN NÚMERO DE APERTURAS(OJO CON EL LABEL DE LOS PUNTOS AZULES)

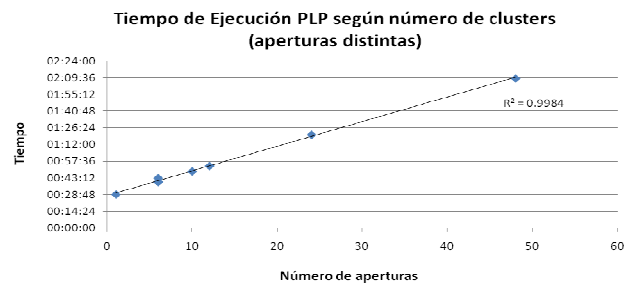


GRÁFICO 3-8: TIEMPOS DE EJECUCIÓN SEGÚN NÚMERO DE CLUSTERS

El Gráfico 3-7 es el mismo gráfico que se ha utilizado a lo largo de este trabajo, en donde se aprecia el tiempo de ejecución en función del número de aperturas. En él puede apreciarse la tendencia lineal que siguen aquellos casos donde el número de aperturas coincide con el número de clusters, así como también aquellos tres puntos que rompen la tendencia lineal (aparte del caso de 3 aperturas). Al analizar este gráfico se ratifica que el tiempo de ejecución del PLP no es siempre lineal con el número de aperturas y dependerá de cuáles aperturas son incluidas en los archivos de entrada. Por otra parte, el Gráfico 3-8 presenta una tendencia lineal entre el tiempo de ejecución y el número de clusters, lo que refuerza la idea de que la respuesta del modelo PLP es distinta ya sea que se trate de aperturas diferentes o repetidas.

Para evaluar el desempeño del programa con respecto a la naturaleza de las aperturas seleccionadas, se ejecutan en forma independiente los problemas lineales que se tienen en las etapas de optimización y de simulación del PLP. A partir de estos resultados, que pueden ser consultados en los anexos, se tiene que los problemas con aperturas replicadas presentan sistemáticamente un tiempo de ejecución menor que aquellos problemas con el mismo número de aperturas, pero en las que todas son distintas. Esto induce a pensar que la complejidad del problema disminuye al utilizar más de una vez la misma apertura, en comparación con utilizar una nueva diferente de la anterior.

Este análisis constituye una herramienta para la decisión final del número de aperturas a utilizar, sobre todo considerando la posibilidad de mantener la noción de equiprobabilidad, sin aumentar los tiempos de ejecución.

3.2.3.2 Silhouette

Como se indicó anteriormente, las agrupaciones propuestas tienen que ver con la experiencia y no con características particulares de la muestra, por esta razón se considera necesario y útil incluir dentro de los criterios de evaluación de las distintas opciones de aperturas disponibles, un índice de validación de clusters que permita identificar la estructura natural de agrupamiento de la muestra.

El índice seleccionado es el de *Silhouette* (en español, silueta) presentado por Rouseeuw en [15] precisamente como herramienta para la interpretación y validación de clusters para determinación del número “correcto” de clusters en los que particionar datos. Este índice $s(i)$ es calculado para cada elemento i de la muestra a partir de la expresión (1).

$$(1) \quad s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Donde:

- $a(i)$ = distancia promedio del elemento i a los elementos del cluster A
- $d(i, C)$ = distancia promedio del elemento i a todos los elementos del cluster C
- $b(i) = \min_{C \neq A} d(i, C)$

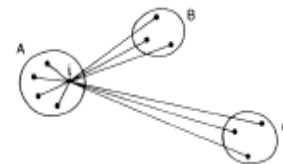


FIGURA 3-3: ELEMENTOS INVOLUCRADOS EN EL CÁLCULO DE $s(i)$

En (1) se evalúa la distancia promedio entre el i -ésimo elemento y los otros elementos del cluster al que pertenece, así como se calculan las distancias del i -ésimo elemento a los elementos de los clusters a los que no pertenece y que luego son promediadas. Por definición, su valor fluctúa entre -1 y 1; mientras más cercano a 1 es su valor significa que los datos se encuentran correctamente agrupados, en cambio, valores negativos indican que probablemente hay valores que necesitan ser reubicados en otros clusters.

Si bien el índice se calcula para cada elemento de la muestra, se puede obtener un valor representativo para la partición realizada al promediar los valores de todos los elementos, obteniéndose $\overline{s(i)}$. Mientras más cercano a 1 sea el $\overline{s(i)}$, el número de clusters seleccionado determina una organización natural de la muestra en esos grupos; mientras que un valor cercano a cero significa que no hay estructura natural de cluster en los datos.

Complementando lo anterior, Rouseeuw propone analizar las figuras que se obtienen al graficar $s(i)$ para toda la muestra, agrupando los elementos en el cluster correspondiente. Este gráfico permite identificar aquellos elementos que se encuentran mal ubicados en su cluster; así como evaluar qué número de clusters es el que produce un valor de *silhouette* mayor, lo que describe como gráficas “más anchas”.

El uso de las dos herramientas: el valor promedio de $s(i)$ para toda la partición y la gráfica de $s(i)$, en forma conjunta permiten identificar el número “correcto” de clusters en los que dividir la muestra.

Finalmente, el cálculo de este algoritmo de validación de clusters se realiza a través de la función *silhouette* de Matlab®. En la Tabla 3-21 se muestran los resultados de este algoritmo para las distintas opciones de clusters; mientras que en la Tabla 3-22, se muestran los resultados para las opciones intermedias entre 6 y 12. A continuación se muestran las gráficas de *silhouette* entre 6 y 12 clusters.

k	$\overline{s(i)}$
3	0.4277
6	0.2668
12	0.2578
24	0.4335
48	1

TABLA 3-21: PROMEDIO DE S(I) PARA OPCIONES INICIALES

k	$\overline{s(i)}$
6	0.1546
7	0.1555
8	0.1547
9	0.1589
10	0.1696
11	0.1661
12	0.1566

TABLA 3-22: PROMEDIO DE S(I) PARA AGRUPAMIENTO ENTRE 6 Y 12 CLUSTERS

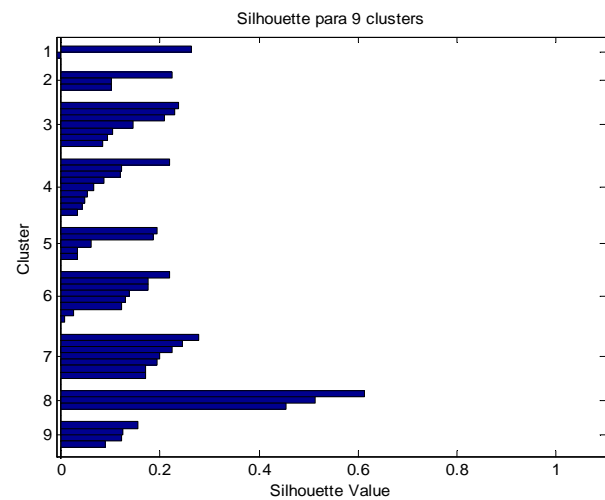
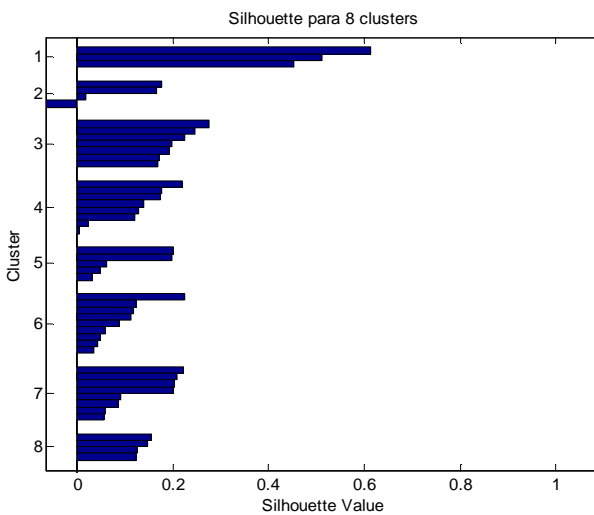
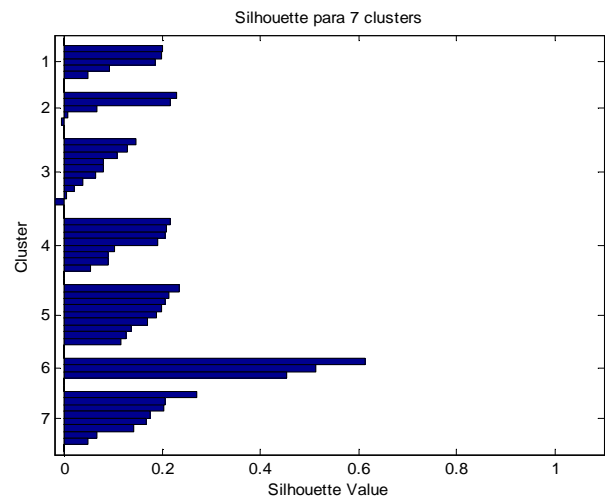
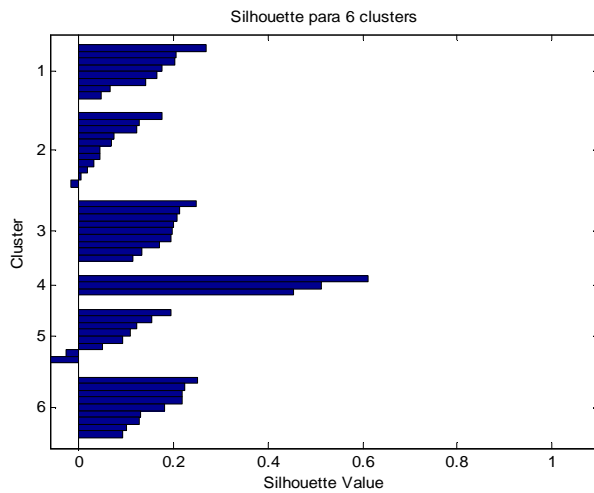


FIGURA 3-4 GRÁFICO DE SILHOUETTE ENTRE 6 Y 9 CLUSTERS

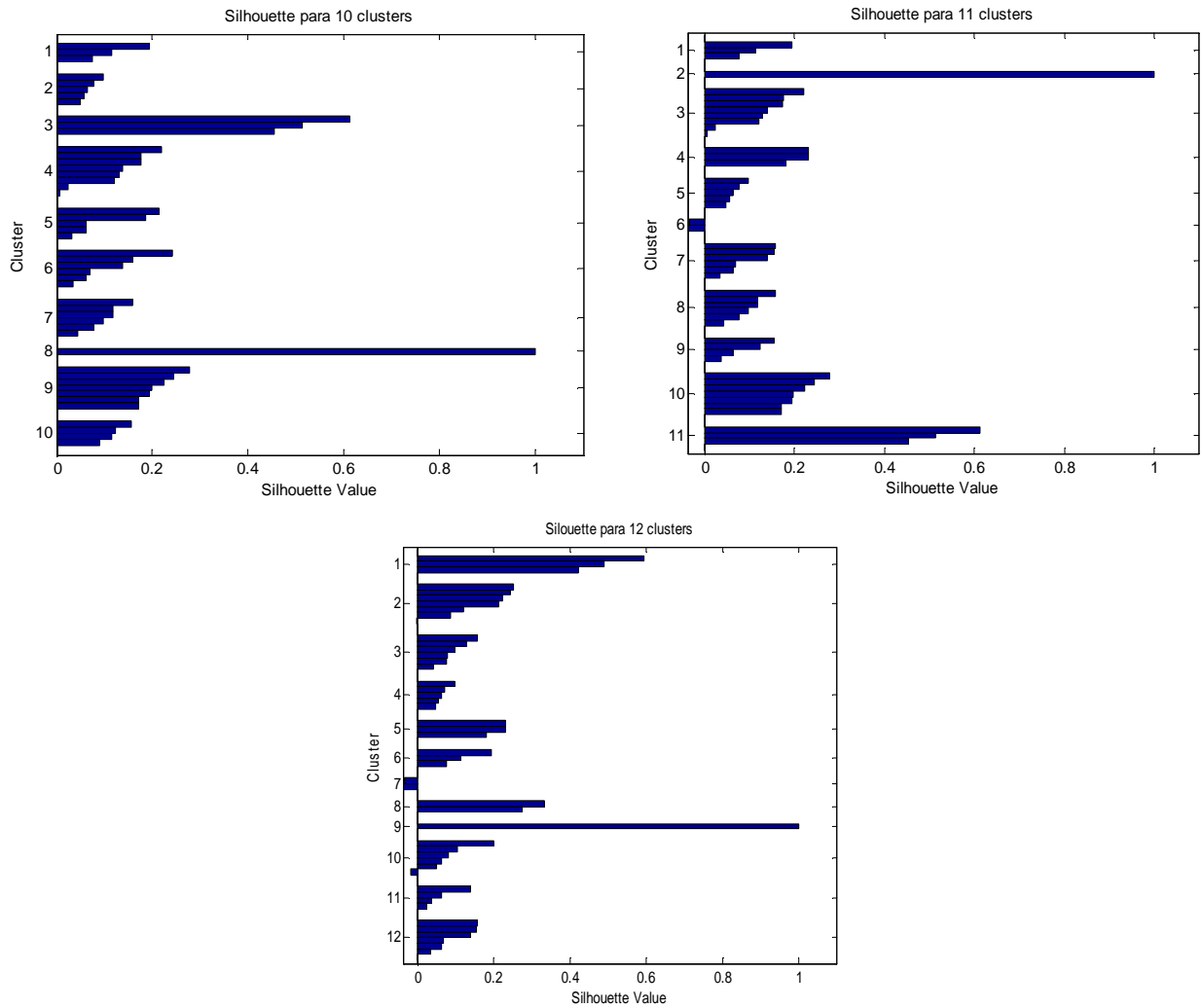


FIGURA 3-5: GRÁFICO DE SILHOUETTE GRÁFICO DE SILHOUETTE ENTRE 6 Y 9 CLUSTERS

El análisis conjunto del valor promedio de silhouette para toda la muestra, así como su gráfica correspondiente permiten discriminar qué asociaciones de clusters representan de mejor manera una agrupación natural de los elementos de aquellas que contienen elementos mal agrupados. Con respecto a las alternativas contempladas originalmente (Tabla 3-21) puede verse que $\overline{s(i)}$ para 3 clusters alcanza un mayor valor que las otras alternativas lo que sugiere que representa una partición natural de la muestra; sin embargo, este caso falla en sus aplicación en el modelo PLP. Como puede verse para la agrupación en 48 clusters, es decir, cuando cada hidrología es considerada como un grupo, $\overline{s(i)}$ alcanza su valor máximo. Esto es producto de la programación de Matlab®, que asigna un valor $s(i)$ igual a 1 a aquellos elementos que conforman por si mismos un grupo; por la misma razón, el caso de 24 aperturas tiene un valor promedio mayor a 6 y 12, pues en él se incluyen 8 grupos con un único elemento. Estos casos (24 y 48), no pueden ser identificados como la mejor representación de la muestra, sino que responden al valor obtenido por el software, es más, Rouseeuw propone asignar un valor $s(i)$ nulo a aquellos grupos de un elemento, para evidenciar que en esos caso no se cuenta con una estructura natural de clusters. En otro sentido, para las opciones de agrupamiento entre 6 y 12 clusters, los valores son comparables entre sí, teniéndose diferencias relativas del 10%.

Evaluando los gráficos de silhouette puede verse que para 6, 8, 11 y 12 clusters, existen elementos que están notoriamente mal asignados a un cluster de acuerdo a estos criterios, lo que no se tiene en los casos de 7, 9 y 10 clusters. De esta manera, estas últimas opciones de partición son las que hay que preferir sobre el rango disponible.

3.2.3.3 Síntesis

Lo mostrado en la presente sección sobre criterios de selección, muestra que al repetir aperturas no se incrementa el tiempo de ejecución y que el índice de validación de clusters propuesto permite acotar el rango disponible a aquellas opciones que no muestran elementos mal asignados en los distintos clusters.

Considerando el tiempo de ejecución, es deseable repetir aperturas de modo de emular el tratamiento de equiprobabilidad con el que actualmente trabaja el modelo PLP, a través de la probabilidad acumulada en los elementos de un cluster; ya que no se incrementan los tiempos con respecto al caso sin repetición. Con respecto al índice de validación, éste permite identificar como válidas las opciones de 7, 9 y 10 clusters.

A fin de aprovechar de mejor manera los criterios planteados se propone la utilización de 7 clusters, definida anteriormente como válida, y en la que todos los grupos tienen una cardinalidad mayor o igual a 3. Como se presentó anteriormente, se busca considerar dicha cardinalidad como la mínima para ponderar cada grupo. A partir de esa consideración se obtiene la Tabla 3-23, que muestra qué aperturas seleccionar y por cuánto ponderar cada una de ellas.

Nº Cluster	Año representativo	Cardinalidad	Ponderación
1	93-94	5	2
2	61-62	5	2
3	94-95	10	3
4	05-06	8	3
5	60-61	9	3
6	98-99	3	1
7	89-90	8	3

TABLA 3-23: OPCIÓN DE 7 CLUSTERS Y 17 APERTURAS

4 CONCLUSIONES

- Con respecto al objetivo general propuesto, se logró desarrollar una metodología sistemática y general para el agrupamiento de series hidrológicas con el propósito de reducir el número de aperturas con las que se alimenta el modelo de coordinación hidrotérmica al interior del CDEC-SIC, Para la selección de aperturas se propuso la agrupación de hidrologías en clusters, para mantener la variabilidad de la muestra con respecto a la original.
- Se logró evaluar el desempeño de la metodología para diversos casos y ajustes en el tratamiento de las series hidrológicas que alimentan al modelo.
- En los casos evaluados se tienen resultados exitosos para la utilización de 24, 12 y 6 aperturas, no así para la opción de 3 aperturas. Más adelante se limitó el rango de búsqueda entre 6 y 12 aperturas, mostrándose que otra opción viable era ponderar al año representativo de cada cluster de acuerdo al número de hidrologías que agrupase, para mantener el tratamiento de equiprobabilidad con el que trabaja el modelo PLP para la estadística histórica.
- Entre los objetivos específicos, se consiguió revisar el estado del arte de las metodologías que resuelven el problema de coordinación hidrotérmica, en particular, la Programación Dinámica Estocástica y la Programación Dinámica Dual Estocástica. En este aspecto, fue de particular interés analizar el modelo PLP que utiliza actualmente el CDEC-SIC.
- Como parte del trabajo de memoria se realizó un estudio de las metodologías que sigue el CDEC-SIC en sus procesos de Planificación de la Operación en lo referente al tratamiento hidrológico, tanto en su dimensión normativa, como de modelo. Lo cual requirió el análisis de las series hidrológicas utilizadas para alimentar el modelo PLP, desde sus características particulares, hasta entender el carácter sistémico del SIC recogido en la matriz de energía afluente.

- Se analizaron las posibilidades de reducción del tiempo de ejecución del modelo PLP al disminuir el número de aperturas, obteniéndose importantes reducciones de tiempo, así como una relación lineal entre el número de aperturas distintas que se seleccionen y el tiempo de ejecución.
- Por otra parte se evaluó la calidad de las alternativas de reducción de aperturas propuestas en esta memoria a través de la revisión de las series de costos marginales en la barra Quillota 220 kV, los costos térmicos del sistema por etapas y el valor del agua en los embalses. Las dos primeras variables evaluadas resultaron ser menos sensibles a la reducción de aperturas que el valor del agua, aunque esta última, tampoco presenta fuertes dispersiones respecto al caso base de 48 hidrologías.
- Como propuesta de estrategia se presenta la utilización de 7 clusters, repitiendo el año representativo del grupo de acuerdo a la cardinalidad de éste; lo que da origen a la utilización de 17 aperturas. Esta recomendación se basa en la evaluación del tiempo de ejecución y de la herramienta silhouette como índice de validación de cluster. Por el tiempo de ejecución, se demostró que el repetir aperturas no repercute negativamente en los tiempos de ejecución, por lo que el tiempo de ejecución de estos casos será considerablemente menor al de 48 aperturas. Mientras que el análisis con silhouette permitió ver que en la partición en 7 clusters no hay elementos que se encuentren mal asignados, lo cual permite seleccionar una división de la muestra que se asemeje a una división natural.
- En el trabajo desarrollado se ha explorado la agrupación de las hidrologías en clusters, dejando abierta otras posibilidades como son el Análisis de Componentes Principales y el de Series Sintéticas. Ambos enfoques pueden ser desarrollados como alternativa al trabajo propuesto, sin embargo, dados los resultados satisfactorios que avalan la aplicación de esta metodología, se proponen desarrollos futuros que trabajen con el software PLP. Ya que el trabajo seguido se enmarcó dentro del código y la interfaz actuales del PLP, sin la capacidad de introducir modificaciones, muchas decisiones debieron tomarse en virtud a la implementación actual sin poder explorar otras posibilidades. Sin perjuicio de lo anterior, como trabajo posterior, se recomienda evaluar nuevas alternativas como la selección de aperturas por cuenca o por embalse o incluso por período (invierno – deshielo) y no de manera agregada; considerando de igual forma que no todos los embalses requieren de la misma cantidad de aperturas.

REFERENCIAS

1. **BROKERING, W., PALMA, R. y VARGAS, L.** *Ñom Lufke (El Rayo Domado) o Los Sistemas Eléctricos de Potencia.* Santiago de Chile : Prentice Hall - Pearson Educación, 2008, págs. 435-444.
2. **SUBSECRETARÍA DE ECONOMÍA, FOMENTO Y RECONSTRUCCIÓN.** Decreto 291/2008. *APRUEBA REGLAMENTO QUE ESTABLECE LA ESTRUCTURA, FUNCIONAMIENTO Y FINANCIAMIENTO DE LOS CENTROS DE DESPACHO ECONÓMICO DE CARGA.*
3. **PEREIRA, M., CAMPODÓNICO, N. and KELMAN, R.** *Application of Stochastic Dual DP and Extension to Hydrothermal Scheduling.* s.l. : PSRI Technical Report, 1999.
4. **TORRES A., R.** *Modelo de Coordinación Hidrotérmica Multietapa y Multiembalse para Sistemas Eléctricos Longitudinales en el Mediano Plazo.* Santiago. : Tesis para optar al grado de Magíster en Ciencias de la Ingeniería, Mención: Ingeniería Eléctrica. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, 2006.
5. **RUDNICK, H. y MOITRE, D.** *Programación Óptima en Mercados de Energía Eléctrica Competitivos al por Mayor.* [ed.] XII Congreso Chileno de Ingeniería Eléctrica Anales Vol I. Temuco : Universidad de la Frontera, 3 al 8 Nov, 1997. págs. 188-195.
6. **OLMEDO, J. C.** Modelo GOL. [En línea] Noviembre de 2001. <http://web.ing.puc.cl/~power/education/GOL%20JColmedo.pdf>.
7. **BIRGE, J. y LOUVEAUX, F.** *Introduction to Stochastic Programming.* New York : Springer series in operations research, 1997.
8. **DIRECCIÓN DE OPERACIÓN CDEC-SIC.** *Informe Final: Puesta en marcha Modelo PLP.* 2003.
9. **BROWN, E., VARGAS X., SALAS J.** *Selección de Modelos Arima Para Series de Tiempo Hidrológicas.* s.l. : Centro de Recursos Hidraulicos, Universidad de Chile., 1983.
10. **PALACIOS, E.** *Modelo para manejo de la incertidumbre hidrológica en la planificación de la operación del SIC.* s.l. : Memoria para optar al Título de Ingeniero Eléctricista. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, 2003.
11. **MINDER, E.** *“Modelación Hidrológica en el Sistema Interconectado Central”*, Tesis para optar al grado de Magíster en Ciencias de la Ingeniería. Santiago de Chile : Pontificia Universidad Católica de Chile, 1997.
12. **WILKS S., D.** *Statistical Methods in the Atmospheric Sciences.* s.l. : International Geophysics Series v.59, 1995.

-
13. **AUGUSTO A, H.** *Método de Clusters para la Predicción Hidrológica en los Problemas de Planificación en el Sistema Interconectado Central.* s.l. : Memoria para optar al Título de Ingeniero Eléctricista. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, 2001.
 14. **BORGES, C. y PINTO, R.** Small Hydro Power Plants Energy Availability Modeling for Generation Reliability Evaluation. s.l. : IEEE TRANSACTIONS ON POWER SYSTEMS, 2008, Vol. 23.
 15. **ROUSEEUW, P.** *Shilbonettes: a graphical aid to the interpretation and validation of clusters analysis.* s.l. : Journal of Computational and Applied Mathematics., 1987.
 16. **GAN, y otros.** *Data Clustering: Theory, Algorithms, and Applications.* Philadelphia : ASA-SIAM Series on Statistics and Applied Probability, 2007.

ANEXO A: ELEMENTOS MATEMÁTICOS

CLUSTERS⁴

ANTECEDENTES GENERALES

El análisis de clusters es un método para crear grupos de objetos, llamados clusters, de manera que los objetos contenidos en un clusters son similares entre sí y diferentes a los objetos presentes en otros clusters. El análisis de clusters es usualmente confundido con la clasificación de objetos, en la cual los datos son asignados a clases predefinidas. En el análisis de clusters, las clases son definidas en el proceso y no se encuentran predefinidas.

La distancia y la similitud juegan un rol importante en el análisis de cluster. En general, distancia y similitud son conceptos recíprocos. Usualmente, las mediciones de similitud y los coeficientes de similitud son usados para describir cuantitativamente cuán parecidos son dos datos de la muestra; mientras mayor sea el coeficiente de similitud, más parecidos son ambos puntos. La medida de la disimilitud y la distancia tienen el enfoque contrario; a mayor distancia, los elementos resultan más disímiles.

Como un problema fundamental de reconocimiento de patrones, un algoritmo de clustering bien diseñado usualmente las siguientes cuatro etapas de diseño: representación de los datos, modelación, optimización y validación. La fase de representación de datos pre-determina qué tipo de estructura de clusters puede ser descubierta en los datos. Basado en lo anterior, la fase de modelación define la noción de clusters y el criterio que separa las estructuras de grupo buscadas, de las desfavorables.

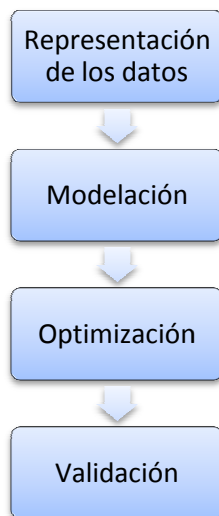


FIGURA ANEXO 1: PROCESO DE CLUSTERING

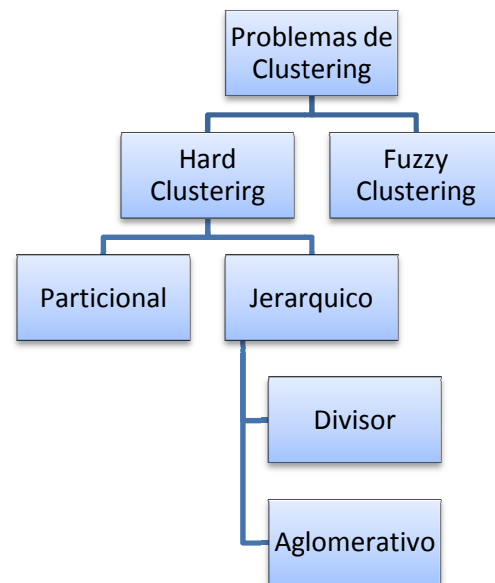


FIGURA ANEXO 2: TIPOS DE CLUSTERING

Generalmente, los problemas de clustering pueden ser divididos en dos categorías: *hard clustering* y *fuzzy clustering*. En la primera categoría, los datos de la muestra pertenecen sólo a un cluster, mientras que en la segunda, los datos pueden pertenecer a dos o más clusters de acuerdo a una probabilidad.

⁴ Para la realización de esta parte del anexo se tomó como referencia [16], por lo que se sugiere para mayor información.

Por otra parte, los algoritmos de clustering convencionales (*hard clustering*) son clasificados en: algoritmos jerárquicos y algoritmos de partición (o no jerárquicos). Hay dos tipos de agrupamiento jerárquico: divisor y aglomerativo. El tipo divisor comienza desde “arriba hacia abajo”, es decir, el algoritmo comienza con un gran cluster conteniendo todos los datos y continúa dividiendo clusters; el aglomerativo procede desde abajo hacia arriba, es decir, el algoritmo comienza con clusters formados por un único elemento y continúa agrupando clusters.

Para gran cantidad de datos, los métodos jerárquicos resultan poco viables, a menos que se incorporen otras técnicas, porque usualmente estos métodos son $O(n^2)$ para requerimientos de espacio en memoria y $O(n^3)$ para tiempos de ejecución en CPU.

DISTANCIA EUCLIDIANA

La selección de la distancia dependerá de la aplicación y la mejor alternativas es usualmente tomada gracias a la combinación de experiencia, habilidad, conocimiento y suerte. Para este trabajo de memoria se seleccionó la distancia euclidiana, la que es probablemente la medida más usual utilizada en datos numéricos. Para dos datos x e y en un espacio de dimensión d , la distancia euclidiana entre ellos se define como:

$$d_{\text{euc}}(x, y) = \left[\sum_{j=1}^d (x_j - y_j)^2 \right]^{\frac{1}{2}}$$

Donde x_j e y_j son los elementos i -ésimos de x e y , respectivamente.

ALGORITMO K-MEANS

El algoritmo convencional de *k-means* fue descrito por primera vez en 1967 por Macqueen. Fue diseñado para datos numéricos, en el que el representativo del cluster se denomina centroide y constituye un tipo de algoritmo particional o no-jerárquico. En este algoritmo, el número k de clusters a formar se considera fijo. Dado un conjunto inicial de k datos (denominados semillas), el método comienza a distribuir los datos restantes al cluster más cercano y luego reasigna los puntos a los clusters de acuerdo a su función de error hasta que esta función de error no cambia significativamente o no se producen re-acomodación de los datos en los clusters.

El algoritmo convencional de k-means se describe brevemente a continuación. Suponiendo, por ejemplo, una muestra D y C_1, C_2, \dots, C_k k clusters disjuntos de D , la función de error está definida como:

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

Donde $\mu(C_i)$ es el centroide del cluster C_i , $d(x, \mu(C_i))$ denota la distancia entre x y $\mu(C_i)$, típicamente la distancia euclidean, pero puede ser otra.

Algoritmo convencional de k-means

Datos de Entrada: Datos originales D , Número de cluster k , Dimensión de los datos d

{ C_i el i -ésimo cluster}

{1. Fase de inicialización}

1: $(C_1, C_2, \dots, C_k) =$ Partición inicial de D

{2. Fase Iterativa}

2: Repetición

3: $d_{ij} =$ distancia entre dato i y cluster j ;

4: $n_i = \arg \min_{1 \leq j \leq k} d_{ij}$;

5: Asignar dato i a cluster n_i ;

6: Recalcular el promedio del cluster para cualquier cluster que haya sido modificado en los pasos anteriores.

7: Se continúa la repetición hasta que no se produzca reubicación de los elementos en los clusters.

8: Se entregan los resultados.

Puede verse que el algoritmo k-means está dividido en dos etapas: la fase de inicialización y la fase iterativa. En la fase de inicialización, el algoritmo trabaja con las semillas elegidas aleatoriamente o entregadas como datos de entrada. En la fase iterativa, el algoritmo calcula la distancia entre cada dato y cada cluster, para asignar el dato al cluster más cercano.

ANEXO B: DESCRIPCIÓN DE ETAPAS

A continuación se muestran las etapas que componen el horizonte de planificación utilizado en los casos del 17/04/2009 y 28/08/2009. Para cada una de los casos, el horizonte de planificación comienza en el día correspondiente y termina con una etapa mensual en marzo de 2010. Las primeras 20 etapas corresponden a las 4 primeras semanas del horizonte de planificación, representándose cada semana por 5 bloques de duración variable. A partir de la etapa 21 se tienen etapas mensuales de un bloque.

Etapa	Tipo de etapa	Año	Mes	Duración [Horas]
1	Bloque 01	2009	Abr	16
2	Bloque 02	2009	Abr	62
3	Bloque 03	2009	Abr	32
4	Bloque 04	2009	Abr	45
5	Bloque 05	2009	Abr	13
6	Bloque 01	2009	Abr	16
7	Bloque 02	2009	Abr	62
8	Bloque 03	2009	Abr	32
9	Bloque 04	2009	Abr	45
10	Bloque 05	2009	Abr	13
11	Bloque 01	2009	May	35
12	Bloque 02	2009	May	39
13	Bloque 03	2009	May	23
14	Bloque 04	2009	May	51
15	Bloque 05	2009	May	20
16	Bloque 01	2009	May	49
17	Bloque 02	2009	May	42
18	Bloque 03	2009	May	32
19	Bloque 04	2009	May	30
20	Bloque 05	2009	May	39
21	Bloque Media	2009	May	384
22	Bloque Media	2009	Jun	720
23	Bloque Media	2009	Jul	744
24	Bloque Media	2009	Ago	744
25	Bloque Media	2009	Sep	720
26	Bloque Media	2009	Oct	744
27	Bloque Media	2009	Nov	720
28	Bloque Media	2009	Dic	744
29	Bloque Media	2010	Ene	744
30	Bloque Media	2010	Feb	672
31	Bloque Media	2010	Mar	744
32	Bloque Media	2010	Abr	720
33	Bloque Media	2010	May	744
34	Bloque Media	2010	Jun	720
35	Bloque Media	2010	Jul	744
36	Bloque Media	2010	Ago	744
37	Bloque Media	2010	Sep	720
38	Bloque Media	2010	Oct	744
39	Bloque Media	2010	Nov	720
40	Bloque Media	2010	Dic	744
41	Bloque Media	2010	Ene	744
42	Bloque Media	2010	Feb	672
43	Bloque Media	2010	Mar	744

TABLA ANEXO 1: DEFINICIÓN DE ETAPAS PARA CASO 17/04/2009

Etapa	Tipo de etapa	Año	Mes	Duración [Horas]
1	Bloque 01	2009	Ago	50
2	Bloque 02	2009	Ago	33
3	Bloque 03	2009	Ago	31
4	Bloque 04	2009	Ago	39
5	Bloque 05	2009	Sep	15
6	Bloque 01	2009	Sep	24
7	Bloque 02	2009	Sep	20
8	Bloque 03	2009	Sep	13
9	Bloque 04	2009	Sep	27
10	Bloque 05	2009	Sep	12
11	Bloque 01	2009	Sep	61
12	Bloque 02	2009	Sep	46
13	Bloque 03	2009	Sep	26
14	Bloque 04	2009	Sep	47
15	Bloque 05	2009	Sep	12
16	Bloque 01	2009	Sep	41
17	Bloque 02	2009	Sep	32
18	Bloque 03	2009	Sep	16
19	Bloque 04	2009	Sep	51
20	Bloque 05	2009	Sep	28
21	Bloque Media	2009	Sep	192
22	Bloque Media	2009	Oct	744
23	Bloque Media	2009	Nov	720
24	Bloque Media	2009	Dic	744
25	Bloque Media	2010	Ene	744
26	Bloque Media	2010	Feb	672
27	Bloque Media	2010	Mar	744
28	Bloque Media	2010	Abr	720
29	Bloque Media	2010	May	744
30	Bloque Media	2010	Jun	720
31	Bloque Media	2010	Jul	744
32	Bloque Media	2010	Ago	744
33	Bloque Media	2010	Sep	720
34	Bloque Media	2010	Oct	744
35	Bloque Media	2010	Nov	720
36	Bloque Media	2010	Dic	744
37	Bloque Media	2010	Ene	744
38	Bloque Media	2010	Feb	672
39	Bloque Media	2010	Mar	744

TABLA ANEXO 2: DEFINICIÓN DE ETAPAS PARA CASO 28/08/2009

ANEXO C: RESULTADOS PLP

COMPARACIÓN DE RESULTADOS

El criterio de evaluación de las series obtenidas consiste en identificar en qué etapas las series obtenidas a partir de la reducción de aperturas se alejan del caso base de 48 aperturas y en qué porcentaje se da esta dispersión. Para esto, se propone definir bandas de tolerancia de 5% en torno al valor base en la etapa t : x'_0 , asignando un color a cada rango para una mejor visualización:

	$0.95 \cdot x'_0 < x' < 1.05 \cdot x'_0$
	$\{0.90 \cdot x'_0 < x' \leq 0.95 \cdot x'_0\} \vee \{1.05 \cdot x'_0 \leq x' < 1.10 \cdot x'_0\}$
	$\{0.85 \cdot x'_0 < x' \leq 0.90 \cdot x'_0\} \vee \{1.10 \cdot x'_0 \leq x' < 1.15 \cdot x'_0\}$
	$\{x' \leq 0.85 \cdot x'_0\} \vee \{1.15 \cdot x'_0 \leq x'\}$

TABLA ANEXO 3: RANGOS DE TOLERANCIA DEFINIDOS

Lo que gráficamente puede ser representado por:

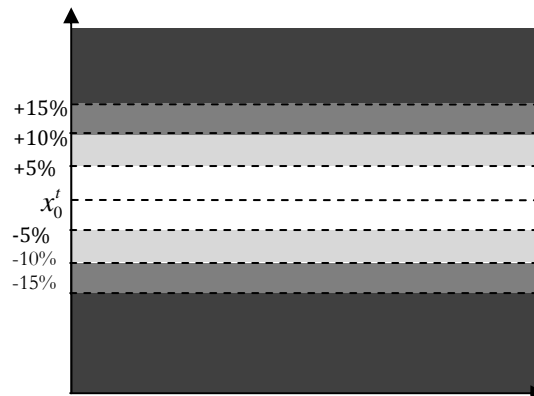


FIGURA ANEXO 3: RANGOS DE TOLERANCIA

Con esto se espera conseguir un análisis cualitativo de la calidad de ajuste que tienen los resultados con un menor número de aperturas al caso base.

En las tablas que se muestran a continuación se sigue la asignación de etapas que se encuentra en el Anexo B.

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	117.79	116.58	118.12	119.17	143.51
2	Bloque 02	113.92	113.69	114.93	115.16	134.08
3	Bloque 03	112.54	111.65	112.20	114.10	120.37
4	Bloque 04	102.00	102.00	102.00	102.00	95.13
5	Bloque 05	92.80	92.80	92.80	92.80	92.80
6	Bloque 01	117.05	115.93	116.60	118.42	132.53
7	Bloque 02	117.18	117.66	117.30	118.43	127.63
8	Bloque 03	113.66	114.42	113.87	114.92	117.78
9	Bloque 04	98.59	98.29	98.74	99.23	99.21
10	Bloque 05	93.48	93.01	93.08	94.13	93.08
11	Bloque 01	113.51	114.07	113.24	114.14	121.41
12	Bloque 02	107.90	107.33	107.37	108.91	115.01
13	Bloque 03	99.32	98.80	99.76	101.32	102.02
14	Bloque 04	93.79	93.05	94.38	94.50	94.02
15	Bloque 05	91.35	90.75	92.24	92.24	91.38
16	Bloque 01	107.77	108.73	107.65	108.88	112.85
17	Bloque 02	110.13	110.66	110.15	109.74	105.09
18	Bloque 03	98.62	98.17	99.01	101.02	101.38
19	Bloque 04	93.68	93.00	94.22	95.10	95.16
20	Bloque 05	92.19	91.40	92.26	92.98	92.57
21	Bloque Media	99.19	99.24	99.31	99.21	100.82
22	Bloque Media	99.71	100.00	99.50	99.08	100.13
23	Bloque Media	92.82	92.41	93.27	93.01	94.11
24	Bloque Media	91.71	91.71	92.89	92.74	92.39
25	Bloque Media	90.24	90.66	90.33	89.83	89.09
26	Bloque Media	88.38	88.12	88.27	88.43	88.29
27	Bloque Media	87.16	87.13	86.83	86.90	86.40
28	Bloque Media	86.01	85.63	86.09	85.95	86.11
29	Bloque Media	101.51	102.18	100.84	100.07	103.05
30	Bloque Media	116.38	114.97	114.33	115.68	110.77
31	Bloque Media	128.31	125.72	125.54	133.53	127.58
32	Bloque Media	130.22	127.68	128.84	140.95	135.89
33	Bloque Media	131.68	133.90	131.36	130.21	132.40
34	Bloque Media	122.63	122.25	123.13	124.05	123.80
35	Bloque Media	86.61	85.82	86.97	88.59	89.43
36	Bloque Media	83.18	82.87	83.05	82.65	82.85
37	Bloque Media	89.93	89.94	89.39	89.66	88.12
38	Bloque Media	89.69	89.50	89.73	89.54	89.40
39	Bloque Media	90.76	90.95	91.27	90.78	90.89
40	Bloque Media	87.28	87.87	87.28	86.09	86.45
41	Bloque Media	83.29	83.28	83.27	83.11	83.18
42	Bloque Media	83.51	83.43	83.35	83.31	82.88
43	Bloque Media	89.34	89.33	89.15	89.18	89.31

TABLA ANEXO 4: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, VALOR ESPERADO EN CASO 17/04/2009

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	107.01	107.00	107.01	107.10	107.10
2	Bloque 02	107.00	107.00	107.00	107.00	107.00
3	Bloque 03	107.00	107.00	107.00	107.00	106.79
4	Bloque 04	103.27	104.47	102.06	102.66	103.67
5	Bloque 05	54.49	51.87	53.50	55.96	50.00
6	Bloque 01	107.01	107.00	107.00	107.10	107.10
7	Bloque 02	106.99	107.00	107.00	107.00	106.98
8	Bloque 03	104.96	105.45	106.01	105.40	103.55
9	Bloque 04	90.45	91.76	90.98	91.27	98.04
10	Bloque 05	55.57	55.56	55.62	56.19	56.99
11	Bloque 01	107.34	107.34	107.29	107.29	105.41
12	Bloque 02	101.89	104.29	102.95	103.10	74.93
13	Bloque 03	88.97	89.19	91.51	89.64	67.45
14	Bloque 04	65.59	65.60	65.66	65.61	62.43
15	Bloque 05	55.40	55.45	55.49	55.51	54.40
16	Bloque 01	92.71	92.67	94.73	93.16	79.56
17	Bloque 02	88.17	88.00	90.32	88.12	72.26
18	Bloque 03	68.93	68.93	68.93	68.93	66.48
19	Bloque 04	59.16	59.25	59.20	59.31	60.44
20	Bloque 05	51.28	50.10	50.49	51.13	53.16
21	Bloque Media	88.04	89.19	88.52	88.66	86.59
22	Bloque Media	83.74	83.53	83.47	83.91	83.30
23	Bloque Media	79.79	79.83	79.62	79.71	76.57
24	Bloque Media	79.46	79.04	78.89	79.73	77.96
25	Bloque Media	73.85	73.87	73.90	74.72	73.31
26	Bloque Media	79.22	78.26	78.83	80.79	78.40
27	Bloque Media	80.73	79.61	80.08	82.57	82.94
28	Bloque Media	82.13	81.57	81.24	83.20	84.44
29	Bloque Media	71.52	73.71	71.54	69.37	72.60
30	Bloque Media	69.20	69.76	69.24	69.35	68.34
31	Bloque Media	67.48	67.84	67.40	67.43	68.05
32	Bloque Media	69.03	68.68	69.77	69.55	69.39
33	Bloque Media	70.49	71.43	70.52	69.86	68.23
34	Bloque Media	64.80	64.93	64.71	64.79	63.93
35	Bloque Media	65.33	65.34	65.22	65.09	64.94
36	Bloque Media	62.23	62.19	62.09	62.05	62.38
37	Bloque Media	66.24	66.36	66.34	66.35	66.68
38	Bloque Media	72.65	72.66	72.60	72.63	72.74
39	Bloque Media	74.84	74.90	74.76	74.71	74.86

TABLA ANEXO 5: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, VALOR ESPERADO EN CASO 28/08/2009

Etapa	k=48	k=24	k=12	k=6	k=3
1-21	116,995.90	115,809.43	118,641.92	122,418.79	129,126.27
22	88,948.47	89,311.78	88,054.69	90,487.46	100,642.53
23	77,790.34	77,109.46	77,826.12	77,472.64	86,034.19
24	73,577.55	73,558.45	76,966.56	76,969.27	78,911.45
25	67,097.37	68,708.62	66,777.55	65,504.37	62,701.15
26	82,408.59	82,232.35	83,255.13	82,819.27	77,972.19
27	71,102.56	71,382.01	71,023.66	70,629.18	69,191.61
28	82,113.59	81,502.31	81,192.66	81,176.32	78,164.72
29	84,358.03	84,688.46	84,141.79	85,818.27	86,790.10
30	97,305.85	95,751.94	97,780.75	100,869.49	97,941.30
31	123,253.65	116,437.89	122,600.52	127,894.40	127,150.55
32	140,972.26	139,436.18	142,333.04	141,368.73	143,709.99
33	124,329.97	128,422.08	122,459.27	120,282.61	121,187.35
34	103,220.77	104,387.16	103,244.48	103,399.28	102,501.76
35	77,441.48	76,444.60	77,308.01	77,921.46	82,561.45
36	74,071.83	72,889.37	74,982.08	74,690.20	73,918.29
37	69,997.60	71,100.11	69,019.37	67,743.11	63,201.57
38	78,389.73	79,247.73	77,723.97	77,610.04	75,242.35
39	67,531.37	67,606.76	67,462.94	68,200.53	66,709.29
40	70,012.25	70,333.91	69,612.64	68,431.73	68,533.30
41	77,264.23	77,544.75	77,208.29	76,817.06	77,101.08
42	89,682.27	89,956.07	89,693.35	89,054.18	88,385.52
43	117,047.66	116,363.68	116,647.78	116,597.96	116,385.29

TABLA ANEXO 6: COSTOS TÉRMICOS POR ETAPA EN [MUS\$], VALOR ESPERADO EN PARA CASO 27/04/2009

Etapa	k=48	k=24	k=12	k=6	k=3
1-21	44,965.63	44,893.86	44,921.57	45,057.40	46,075.81
22	52,453.91	52,448.86	52,657.40	53,272.31	51,539.33
23	48,110.48	47,698.32	47,926.07	48,061.94	47,586.22
24	49,567.10	49,239.40	49,418.73	50,170.27	52,561.63
25	63,213.90	62,842.59	62,976.27	63,395.96	63,839.26
26	64,814.84	63,907.86	64,803.14	67,739.30	67,291.16
27	77,649.78	76,851.42	76,847.54	79,102.79	80,302.74
28	79,038.10	77,989.35	78,015.31	79,220.69	78,416.61
29	59,799.85	59,770.25	59,949.38	55,733.66	62,425.23
30	49,597.83	50,046.19	49,391.21	50,137.31	51,180.41
31	46,295.96	46,923.11	46,789.52	46,079.34	46,877.05
32	54,396.27	53,321.43	54,574.02	54,822.59	54,196.92
33	51,456.84	52,879.23	51,651.70	50,498.04	48,731.41
34	53,058.76	53,510.05	53,352.43	53,227.05	52,072.28
35	45,426.48	45,666.63	45,455.54	45,617.61	46,099.63
36	51,449.34	51,557.02	51,429.17	51,393.24	51,287.73
37	60,221.56	60,267.49	60,301.15	60,451.90	60,457.01
38	65,732.40	65,785.84	65,817.32	65,537.51	65,491.41
39	78,774.77	78,679.64	78,569.24	78,276.24	78,757.68

TABLA ANEXO 7: COSTOS TÉRMICOS POR ETAPA EN [MUS\$], VALOR ESPERADO EN PARA CASO 28/08/2009

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	117.79	116.58	118.12	119.17	143.51
2	Bloque 02	113.92	113.69	114.93	115.16	134.08
3	Bloque 03	112.54	111.65	112.20	114.10	120.37
4	Bloque 04	102.00	102.00	102.00	102.00	95.13
5	Bloque 05	92.80	92.80	92.80	92.80	92.80
6	Bloque 01	117.41	115.87	116.68	118.16	131.88
7	Bloque 02	113.15	112.97	113.51	114.78	127.42
8	Bloque 03	109.37	108.96	109.54	111.37	107.19
9	Bloque 04	95.13	94.75	94.75	95.91	94.75
10	Bloque 05	92.80	92.80	92.80	92.80	92.80
11	Bloque 01	108.39	107.96	108.20	109.04	120.19
12	Bloque 02	102.87	102.00	103.91	105.50	112.34
13	Bloque 03	96.02	95.42	96.76	100.10	102.00
14	Bloque 04	93.20	92.80	94.25	94.75	92.80
15	Bloque 05	90.60	90.60	90.60	90.60	90.60
16	Bloque 01	102.00	102.00	102.00	102.75	110.08
17	Bloque 02	99.29	98.23	99.38	102.00	100.67
18	Bloque 03	94.25	93.82	95.00	98.18	98.60
19	Bloque 04	92.80	91.80	92.80	95.00	93.09
20	Bloque 05	90.60	90.60	90.60	91.25	91.65
21	Bloque Media	90.50	90.50	90.50	90.60	95.62
22	Bloque Media	90.50	90.50	90.50	90.50	90.50
23	Bloque Media	78.50	78.50	78.50	78.50	88.16
24	Bloque Media	78.50	78.50	78.50	78.50	78.50
25	Bloque Media	78.50	78.50	78.50	78.50	78.50
26	Bloque Media	87.50	87.50	87.50	87.50	88.79
27	Bloque Media	87.50	87.50	87.50	87.50	87.50
28	Bloque Media	86.05	78.92	85.09	84.60	78.50
29	Bloque Media	87.50	87.50	87.50	87.50	88.63
30	Bloque Media	100.70	100.70	100.70	100.70	100.70
31	Bloque Media	100.70	100.70	100.88	103.01	126.30
32	Bloque Media	106.42	101.05	102.78	122.37	125.36
33	Bloque Media	100.70	100.70	100.70	100.70	100.70
34	Bloque Media	100.70	100.70	100.70	100.70	100.70
35	Bloque Media	61.27	68.07	66.83	87.50	72.01
36	Bloque Media	60.96	60.77	60.86	60.82	60.58
37	Bloque Media	87.50	87.50	87.50	87.50	87.50
38	Bloque Media	87.50	87.50	87.50	87.50	87.50
39	Bloque Media	87.50	87.50	87.50	87.50	87.50
40	Bloque Media	87.50	87.50	87.50	87.50	87.50
41	Bloque Media	87.50	87.50	87.50	87.50	87.50
42	Bloque Media	87.50	87.50	87.50	87.50	87.50
43	Bloque Media	87.50	87.50	87.50	87.50	87.50

TABLA ANEXO 8: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO HÚMEDO (86-87)
EN CASO 17/04/2009

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	117.79	116.58	118.12	119.17	143.51
2	Bloque 02	113.92	113.69	114.93	115.16	134.08
3	Bloque 03	112.54	111.65	112.20	114.10	120.37
4	Bloque 04	102	102.00	102.00	102.00	95.13
5	Bloque 05	92.8	92.80	92.80	92.80	92.80
6	Bloque 01	115.85	115.02	116.66	116.00	131.71
7	Bloque 02	113.15	112.96	113.50	115.01	127.43
8	Bloque 03	109.38	108.97	109.53	111.36	116.20
9	Bloque 04	95.43	94.75	94.75	95.91	94.75
10	Bloque 05	92.8	92.80	92.80	92.80	92.80
11	Bloque 01	102	102.00	102.21	103.97	114.55
12	Bloque 02	94.25	94.25	98.49	101.80	101.84
13	Bloque 03	92.8	92.80	92.80	93.82	92.80
14	Bloque 04	90.6	90.60	90.60	90.60	91.32
15	Bloque 05	90.5	90.50	90.50	90.50	90.50
16	Bloque 01	100.76	97.06	97.28	100.56	105.73
17	Bloque 02	95.37	93.69	94.52	97.93	100.41
18	Bloque 03	92.8	92.80	92.80	95.42	92.80
19	Bloque 04	90.6	90.60	90.60	90.60	91.11
20	Bloque 05	90.5	90.50	90.50	90.50	90.60
21	Bloque Media	90.5	90.50	90.50	90.50	90.50
22	Bloque Media	90.5	90.50	90.50	90.50	90.50
23	Bloque Media	90.5	90.50	90.50	79.68	90.50
24	Bloque Media	78.5	78.50	78.50	78.50	78.50
25	Bloque Media	78.5	78.50	78.50	78.50	78.50
26	Bloque Media	90.5	89.58	90.50	90.50	90.50
27	Bloque Media	90.5	87.98	90.50	90.50	90.50
28	Bloque Media	90.5	90.50	90.50	90.50	90.50
29	Bloque Media	100.7	100.70	100.70	100.70	100.70
30	Bloque Media	126.06	100.70	100.70	126.06	100.70
31	Bloque Media	148.72	151.72	140.67	164.05	100.70
32	Bloque Media	173.41	169.46	166.44	164.70	100.70
33	Bloque Media	175.7	175.70	175.70	175.70	160.69
34	Bloque Media	90.72	87.50	88.20	100.70	100.70
35	Bloque Media	59.6	59.86	59.79	59.86	
36	Bloque Media	60.66	60.49	60.83	60.82	60.94
37	Bloque Media	57.12	57.30	57.30	57.12	57.12
38	Bloque Media	87.5	87.50	87.50	87.50	87.50
39	Bloque Media	87.5	87.50	87.50	87.50	87.50
40	Bloque Media	57.84	57.84	57.84	57.84	57.84
41	Bloque Media	62.86	62.86	62.86	62.86	62.86
42	Bloque Media	74.68	74.68	74.68	74.68	74.68
43	Bloque Media	82.36	82.35	82.36	82.36	82.36

TABLA ANEXO 9: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO MEDIO (81-82) EN CASO 17/04/2009

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	117.79	116.58	118.12	119.17	143.51
2	Bloque 02	113.92	113.69	114.93	115.16	134.08
3	Bloque 03	112.54	111.65	112.20	114.10	120.37
4	Bloque 04	102.00	102.00	102.00	102.00	95.13
5	Bloque 05	92.80	92.80	92.80	92.80	92.80
6	Bloque 01	117.29	116.26	116.66	119.24	132.53
7	Bloque 02	117.79	117.05	117.41	119.04	127.80
8	Bloque 03	111.13	110.78	111.22	113.27	119.01
9	Bloque 04	98.04	97.60	99.83	101.08	102.00
10	Bloque 05	94.74	92.80	92.80	94.74	92.80
11	Bloque 01	109.38	120.14	109.06	111.55	122.53
12	Bloque 02	106.38	112.19	106.01	108.25	119.52
13	Bloque 03	102.00	102.00	102.00	102.00	102.00
14	Bloque 04	94.25	92.81	94.75	94.75	95.00
15	Bloque 05	92.80	90.60	92.80	92.80	92.80
16	Bloque 01	107.67	112.26	106.91	108.48	115.48
17	Bloque 02	136.38	136.38	133.19	129.90	105.21
18	Bloque 03	102.00	102.00	102.00	104.38	102.00
19	Bloque 04	95.00	93.91	95.16	95.42	95.42
20	Bloque 05	92.80	92.80	92.80	92.80	92.80
21	Bloque Media	118.93	118.93	118.93	118.33	105.34
22	Bloque Media	142.30	142.30	142.30	142.06	142.30
23	Bloque Media	92.99	92.99	92.99	92.99	98.34
24	Bloque Media	131.31	132.32	135.50	135.50	134.43
25	Bloque Media	127.90	127.75	127.75	127.75	123.19
26	Bloque Media	90.60	90.60	90.60	90.60	101.24
27	Bloque Media	90.50	90.50	90.50	90.50	90.50
28	Bloque Media	90.50	90.50	90.50	90.50	90.50
29	Bloque Media	118.58	118.58	118.89	118.58	119.18
30	Bloque Media	141.33	143.04	138.39	145.18	129.45
31	Bloque Media	149.00	146.82	148.14	149.78	160.62
32	Bloque Media	165.76	164.70	165.76	164.70	166.21
33	Bloque Media	175.70	175.70	175.70	175.70	175.70
34	Bloque Media	158.11	158.11	158.11	158.11	158.11
35	Bloque Media	100.70	100.70	100.70	100.70	100.70
36	Bloque Media	87.50	87.50	87.50	87.50	87.50
37	Bloque Media	100.70	100.70	100.70	100.70	100.70
38	Bloque Media	87.50	87.50	87.50	87.50	87.50
39	Bloque Media	87.50	87.50	87.50	87.50	87.50
40	Bloque Media	87.50	87.50	87.50	87.50	87.50
41	Bloque Media	87.50	87.50	87.50	87.50	87.50
42	Bloque Media	87.50	87.50	87.50	87.50	87.50
43	Bloque Media	87.50	87.50	87.50	87.50	87.50

TABLA ANEXO 10: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO SECO (07-08) EN CASO 17/04/2009

Caso Agosto

A diferencia del caso de abril, para agosto el horizonte de planificación sólo considera el segundo semestre del año hidrológico seleccionado, por lo que la información contenida en la matriz de energía afluente no es suficiente para determinar si el segundo semestre a considerar dentro de la simulación corresponde a una condición hidrológica húmeda, media o seca. Por esto, se recalcula la energía afluente tomando en cuenta sólo aquellas semanas comprendidas entre el 28 de agosto y el final del año hidrológico y, a partir de estos resultados, se determinan los años con probabilidad de excedencia 20%, 50% y 90% como años: húmedo, medio y seco, respectivamente.

Año Hidrológico	Probabilidad de excedencia	Condición
97-98	20.41%	Húmedo
93-94	51.02%	Medio
07-08	89.80%	Seco

TABLA ANEXO 11: CONDICIONES HIDROLÓGICAS SELECCIONADAS ACUERDO A SU PROBABILIDAD DE EXCEDENCIA

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	107.01	107.00	107.01	107.10	107.10
2	Bloque 02	107.00	107.00	107.00	107.00	107.00
3	Bloque 03	107.00	107.00	107.00	107.00	106.79
4	Bloque 04	103.27	104.47	102.06	102.66	103.67
5	Bloque 05	54.49	51.87	53.50	55.96	50.00
6	Bloque 01	107.00	107.00	107.00	107.10	107.00
7	Bloque 02	107.00	107.00	107.00	107.00	107.00
8	Bloque 03	103.18	103.43	104.43	103.66	103.43
9	Bloque 04	61.00	60.98	61.02	61.02	67.84
10	Bloque 05	55.35	55.34	55.37	55.92	41.43
11	Bloque 01	107.00	107.00	107.00	107.00	103.83
12	Bloque 02	90.43	92.63	95.24	91.90	68.40
13	Bloque 03	83.90	85.76	88.29	84.56	62.91
14	Bloque 04	60.97	60.95	60.98	61.01	55.38
15	Bloque 05	54.51	54.51	54.52	54.55	40.09
16	Bloque 01	83.41	85.57	87.76	85.31	68.40
17	Bloque 02	79.54	80.42	82.24	74.90	65.82
18	Bloque 03	68.40	68.40	68.40	68.40	62.35
19	Bloque 04	48.82	53.85	49.30	52.17	55.72
20	Bloque 05	47.48	46.09	46.82	47.31	38.82
21	Bloque Media	68.40	68.40	68.40	68.40	68.40
22	Bloque Media	68.40	68.40	68.40	68.40	68.40
23	Bloque Media	68.40	68.40	68.40	68.40	68.40
24	Bloque Media	68.40	68.40	68.40	68.40	68.40
25	Bloque Media	59.40	59.40	59.40	59.40	59.40
26	Bloque Media	66.20	66.20	66.20	66.25	66.20
27	Bloque Media	67.71	66.20	67.05	68.21	67.83
28	Bloque Media	67.79	66.47	66.44	69.47	70.97
29	Bloque Media	66.20	66.20	66.20	64.00	66.20
30	Bloque Media	98.50	98.50	98.50	98.50	74.14
31	Bloque Media	104.60	104.60	104.60	104.60	104.60
32	Bloque Media	104.60	104.60	104.60	104.60	104.60
33	Bloque Media	104.60	104.60	104.60	104.60	104.60
34	Bloque Media	104.60	104.60	104.60	104.60	104.60
35	Bloque Media	104.60	104.60	104.60	104.60	104.60
36	Bloque Media	98.50	98.50	98.50	98.50	98.50
37	Bloque Media	110.60	110.60	110.60	110.60	110.60
38	Bloque Media	110.60	110.60	110.60	110.60	110.60
39	Bloque Media	110.60	110.60	110.60	110.60	110.60

TABLA ANEXO 12: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO HÚMEDO (97-98)
EN CASO 28/08/2009

Etapa	Tipo de etapa	k=48	k=24	k=12	k=6	k=3
1	Bloque 01	107.01	107.00	107.01	107.10	107.10
2	Bloque 02	107.00	107.00	107.00	107.00	107.00
3	Bloque 03	107.00	107.00	107.00	107.00	106.79
4	Bloque 04	103.27	104.47	102.06	102.66	103.67
5	Bloque 05	54.49	51.87	53.50	55.96	50.00
6	Bloque 01	107.00	107.00	107.00	107.10	107.00
7	Bloque 02	107.00	107.00	107.00	107.00	107.00
8	Bloque 03	105.99	106.85	107.00	107.00	103.33
9	Bloque 04	100.57	104.35	101.49	101.48	101.75
10	Bloque 05	55.69	55.67	55.69	56.10	55.63
11	Bloque 01	107.00	107.00	107.00	107.00	105.69
12	Bloque 02	107.00	107.00	107.00	107.00	68.40
13	Bloque 03	87.73	87.86	90.55	88.43	64.96
14	Bloque 04	68.40	68.40	68.40	68.40	62.65
15	Bloque 05	55.66	55.64	55.67	55.70	55.63
16	Bloque 01	90.91	89.96	92.95	90.61	68.40
17	Bloque 02	83.63	83.92	86.47	83.72	68.40
18	Bloque 03	68.40	68.40	68.40	68.40	65.92
19	Bloque 04	60.05	60.04	60.06	60.08	60.62
20	Bloque 05	49.37	48.55	49.39	50.38	55.36
21	Bloque Media	102.13	102.13	102.13	102.13	101.54
22	Bloque Media	98.47	98.47	98.47	98.47	98.47
23	Bloque Media	68.40	68.40	68.40	68.40	68.40
24	Bloque Media	68.40	68.40	68.40	68.40	68.40
25	Bloque Media	66.20	66.20	66.20	66.20	66.20
26	Bloque Media	68.20	67.84	67.81	72.59	71.94
27	Bloque Media	71.26	70.23	72.33	74.75	75.01
28	Bloque Media	73.51	70.81	71.36	74.04	76.35
29	Bloque Media	66.20	66.20	66.20	66.20	66.20
30	Bloque Media	59.43	62.15	60.70	62.29	63.11
31	Bloque Media	59.11	58.04	59.14	57.48	57.35
32	Bloque Media	59.40	59.40	59.40	59.40	59.40
33	Bloque Media	59.40	59.40	59.40	59.40	59.40
34	Bloque Media	59.40	59.40	59.40	59.40	59.40
35	Bloque Media	61.67	61.90	61.67	61.36	60.10
36	Bloque Media	59.40	59.40	59.40	59.40	59.40
37	Bloque Media	66.30	66.30	66.30	66.30	66.30
38	Bloque Media	73.45	73.14	73.19	73.90	73.45
39	Bloque Media	73.90	73.90	73.90	73.90	73.90

TABLA ANEXO 13: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO MEDIO (93-94)
EN CASO 28/08/2009

Etapa	k=48	k=24	k=12	k=6	k=3
1	107.01	107.00	107.01	107.10	107.10
2	107.00	107.00	107.00	107.00	107.00
3	107.00	107.00	107.00	107.00	106.79
4	103.27	104.47	102.06	102.66	103.67
5	54.49	51.87	53.50	55.96	50.00
6	107.00	107.00	107.00	107.10	107.00
7	107.00	107.00	107.00	107.00	107.00
8	105.17	106.69	105.74	107.00	103.46
9	101.85	104.39	101.54	103.72	102.63
10	55.64	55.63	55.66	56.10	57.28
11	107.00	107.00	107.00	107.00	107.00
12	107.00	107.00	107.00	107.00	77.94
13	87.69	87.83	90.57	88.42	67.01
14	68.40	68.40	68.40	68.40	62.94
15	55.62	55.60	55.64	55.66	55.46
16	106.02	105.88	106.11	106.31	107.00
17	104.62	104.57	104.71	104.87	75.03
18	68.40	68.40	68.40	68.40	67.74
19	60.01	60.01	60.03	60.04	60.68
20	53.78	53.75	52.19	53.80	53.67
21	104.21	104.11	104.15	104.29	107.00
22	107.10	107.10	107.10	107.10	107.10
23	107.00	107.00	107.00	107.00	68.40
24	107.00	107.00	107.00	107.00	107.10
25	104.60	104.60	104.60	104.60	104.60
26	104.60	104.60	104.60	104.60	104.60
27	104.60	104.60	104.80	104.60	104.80
28	104.60	104.60	104.60	104.60	98.50
29	104.60	104.60	104.60	104.60	104.60
30	66.20	66.20	66.20	66.20	70.36
31	66.20	66.20	66.20	66.20	73.37
32	98.50	98.50	98.50	98.50	98.50
33	104.60	104.60	104.60	104.60	104.60
34	59.40	59.40	59.40	61.34	59.53
35	61.92	61.95	61.92	62.09	61.91
36	59.40	59.40	59.40	59.45	59.40
37	66.30	66.30	66.30	66.30	66.30
38	73.90	73.90	73.90	73.90	73.90
39	73.90	73.90	73.90	73.90	73.90

TABLA ANEXO 14: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, AÑO SECO (07-08) EN CASO 28/08/2009

Etapa	Tipo de etapa	k=48	k=17	k=12	k=10	k=6
1	Bloque 01	117.79	120.81	118.12	117.58	119.17
2	Bloque 02	113.92	115.56	114.93	113.69	115.16
3	Bloque 03	112.54	113.76	112.20	112.83	114.10
4	Bloque 04	102.00	102.00	102.00	102.00	102.00
5	Bloque 05	92.80	92.80	92.80	92.80	92.80
6	Bloque 01	117.05	118.10	116.60	116.73	118.42
7	Bloque 02	117.18	118.13	117.30	118.61	118.43
8	Bloque 03	113.66	114.94	113.87	114.64	114.92
9	Bloque 04	98.59	98.83	98.74	98.80	99.23
10	Bloque 05	93.48	93.23	93.08	93.08	94.13
11	Bloque 01	113.51	114.05	113.24	114.03	114.14
12	Bloque 02	107.90	107.96	107.37	107.86	108.91
13	Bloque 03	99.32	100.14	99.76	100.28	101.32
14	Bloque 04	93.79	94.51	94.38	94.53	94.50
15	Bloque 05	91.35	92.24	92.24	92.20	92.24
16	Bloque 01	107.77	108.39	107.65	107.86	108.88
17	Bloque 02	110.13	109.33	110.15	109.74	109.74
18	Bloque 03	98.62	99.83	99.01	99.45	101.02
19	Bloque 04	93.68	94.77	94.22	94.38	95.10
20	Bloque 05	92.19	92.76	92.26	92.47	92.98
21	Bloque Media	99.19	98.70	99.31	99.53	99.21
22	Bloque Media	99.71	99.45	99.50	99.46	99.08
23	Bloque Media	92.82	92.44	93.27	93.22	93.01
24	Bloque Media	91.71	91.65	92.89	92.12	92.74
25	Bloque Media	90.24	90.60	90.33	90.37	89.83
26	Bloque Media	88.38	88.29	88.27	88.30	88.43
27	Bloque Media	87.16	86.95	86.83	87.06	86.90
28	Bloque Media	86.01	85.59	86.09	86.10	85.95
29	Bloque Media	101.51	102.22	100.84	100.64	100.07
30	Bloque Media	116.38	115.77	114.33	115.33	115.68
31	Bloque Media	128.31	133.53	125.54	131.61	133.53
32	Bloque Media	130.22	138.59	128.84	136.35	140.95
33	Bloque Media	131.68	129.78	131.36	130.74	130.21
34	Bloque Media	122.63	123.16	123.13	123.09	124.05
35	Bloque Media	86.61	86.22	86.97	87.58	88.59
36	Bloque Media	83.18	82.77	83.05	83.11	82.65
37	Bloque Media	89.93	89.94	89.39	89.39	89.66
38	Bloque Media	89.69	89.50	89.73	89.73	89.54
39	Bloque Media	90.76	90.49	91.27	90.98	90.78
40	Bloque Media	87.28	86.21	87.28	86.64	86.09
41	Bloque Media	83.29	83.30	83.27	83.27	83.11
42	Bloque Media	83.51	83.59	83.35	83.49	83.31
43	Bloque Media	89.34	89.23	89.15	89.23	89.18

TABLA ANEXO 15: COSTOS MARGINALES [US\$/MWH] SEGÚN NÚMERO DE CLUSTERS, VALOR ESPERADO PARA EL CASO 17/04/2009

Etapa	Tipo de etapa	k=48	k=17	k=12	k=10	k=6
1	Bloque 01	107.01	107.00	107.01	107.08	107.10
2	Bloque 02	107.00	107.00	107.00	107.00	107.00
3	Bloque 03	107.00	106.62	107.00	107.00	107.00
4	Bloque 04	103.27	103.74	102.06	104.69	102.66
5	Bloque 05	54.49	51.40	53.50	53.46	55.96
6	Bloque 01	107.01	107.01	107.00	107.05	107.10
7	Bloque 02	106.99	106.99	107.00	107.00	107.00
8	Bloque 03	104.96	105.20	106.01	105.51	105.40
9	Bloque 04	90.45	90.85	90.98	91.57	91.27
10	Bloque 05	55.57	55.59	55.62	55.59	56.19
11	Bloque 01	107.34	107.45	107.29	107.29	107.29
12	Bloque 02	101.89	104.70	102.95	104.20	103.10
13	Bloque 03	88.97	88.33	91.51	90.36	89.64
14	Bloque 04	65.59	65.56	65.66	65.63	65.61
15	Bloque 05	55.40	55.44	55.49	55.47	55.51
16	Bloque 01	92.71	92.36	94.73	93.90	93.16
17	Bloque 02	88.17	88.02	90.32	89.15	88.12
18	Bloque 03	68.93	68.93	68.93	68.93	68.93
19	Bloque 04	59.16	59.33	59.20	59.24	59.31
20	Bloque 05	51.28	50.46	50.49	50.66	51.13
21	Bloque Media	88.04	88.94	88.52	88.45	88.66
22	Bloque Media	83.74	84.10	83.47	83.74	83.91
23	Bloque Media	79.79	80.11	79.62	79.81	79.71
24	Bloque Media	79.46	79.53	78.89	79.21	79.73
25	Bloque Media	73.85	74.27	73.90	74.00	74.72
26	Bloque Media	79.22	80.83	78.83	80.02	80.79
27	Bloque Media	80.73	82.44	80.08	81.24	82.57
28	Bloque Media	82.13	82.82	81.24	82.35	83.20
29	Bloque Media	71.52	69.27	71.54	70.92	69.37
30	Bloque Media	69.20	69.28	69.24	69.11	69.35
31	Bloque Media	67.48	67.48	67.40	67.31	67.43
32	Bloque Media	69.03	68.83	69.77	69.45	69.55
33	Bloque Media	70.49	71.09	70.52	70.51	69.86
34	Bloque Media	64.80	64.89	64.71	64.76	64.79
35	Bloque Media	65.33	65.30	65.22	65.21	65.09
36	Bloque Media	62.23	62.13	62.09	62.10	62.05
37	Bloque Media	66.24	66.63	66.34	66.28	66.35
38	Bloque Media	72.65	72.67	72.60	72.62	72.63
39	Bloque Media	74.84	74.74	74.76	74.78	74.71

TABLA ANEXO 16: COSTOS MARGINALES SEGÚN NÚMERO DE CLUSTERS, VALOR ESPERADO PARA EL CASO
17/04/2009

Etapa	k=48	k=17	k=12	k=10	k=6
1 - 21	116,995.90	118,625.79	118,641.92	120,494.04	122,418.79
22	88,948.47	88,281.05	88,054.69	87,598.55	90,487.46
23	77,790.34	77,183.70	77,826.12	77,500.65	77,472.64
24	73,577.55	72,957.60	76,966.56	76,029.67	76,969.27
25	67,097.37	67,939.53	66,777.55	67,333.92	65,504.37
26	82,408.59	81,982.38	83,255.13	82,801.73	82,819.27
27	71,102.56	71,576.80	71,023.66	71,262.54	70,629.18
28	82,113.59	81,598.88	81,192.66	81,457.14	81,176.32
29	84,358.03	85,733.54	84,141.79	85,432.39	85,818.27
30	97,305.85	101,337.32	97,780.75	99,391.98	100,869.49
31	123,253.65	127,777.40	122,600.52	124,537.49	127,894.40
32	140,972.26	141,281.73	142,333.04	142,904.71	141,368.73
33	124,329.97	118,852.74	122,459.27	120,439.37	120,282.61
34	103,220.77	102,752.89	103,244.48	102,201.96	103,399.28
35	77,441.48	76,435.81	77,308.01	76,754.89	77,921.46
36	74,071.83	74,047.31	74,982.08	74,173.83	74,690.20
37	69,997.60	69,965.43	69,019.37	68,884.83	67,743.11
38	78,389.73	78,519.26	77,723.97	78,784.44	77,610.04
39	67,531.37	67,519.96	67,462.94	68,032.34	68,200.53
40	70,012.25	69,279.93	69,612.64	68,385.57	68,431.73
41	77,264.23	77,686.66	77,208.29	77,129.44	76,817.06
42	89,682.27	89,975.44	89,693.35	89,248.14	89,054.18
43	117,047.66	116,815.54	116,647.78	116,658.17	116,597.96

TABLA ANEXO 17: COSTOS TÉRMICOS POR ETAPA [MUS\$] PARA VALOR ESPERADO EN CASO 17/04/2009

Etapa	k =48	k = 17	k =12	k =10	k =6
1 - 21	44,965.63	44,861.11	44,921.57	44,892.96	45,057.40
22	52,453.91	53,241.74	52,657.40	52,887.64	53,272.31
23	48,110.48	48,179.08	47,926.07	47,895.50	48,061.94
24	49,567.10	49,667.65	49,418.73	49,455.42	50,170.27
25	63,213.90	63,539.95	62,976.27	63,481.50	63,395.96
26	64,814.84	67,051.55	64,803.14	66,114.22	67,739.30
27	77,649.78	79,210.96	76,847.54	78,478.07	79,102.79
28	79,038.10	78,282.56	78,015.31	78,652.16	79,220.69
29	59,799.85	54,385.75	59,949.38	58,551.02	55,733.66
30	49,597.83	49,976.37	49,391.21	49,012.33	50,137.31
31	46,295.96	45,493.07	46,789.52	46,385.97	46,079.34
32	54,396.27	53,757.84	54,574.02	54,364.39	54,822.59
33	51,456.84	52,056.77	51,651.70	51,425.57	50,498.04
34	53,058.76	53,501.16	53,352.43	53,324.00	53,227.05
35	45,426.48	45,598.35	45,455.54	45,727.52	45,617.61
36	51,449.34	51,463.21	51,429.17	51,472.90	51,393.24
37	60,221.56	60,528.35	60,301.15	60,471.79	60,451.90
38	65,732.40	65,824.43	65,817.32	65,771.83	65,537.51
39	78,774.77	78,506.53	78,569.24	78,293.65	78,276.24

TABLA ANEXO 18: COSTOS TÉRMICOS POR ETAPA [MUS\$] PARA VALOR ESPERADO EN CASO 28/08/2009

PROBLEMAS LINEALES EN CPLEX

A continuación se presentan los tiempos de ejecución de los problemas lineales que resultan para cada iteración, fase, etapa, simulación y apertura, solucionados mediante CPLEX. Como se explica en 3.2.3.1, el objetivo es evaluar si el tiempo de ejecución de problemas definidos con aperturas replicadas es sistemáticamente menor que el tiempo de problemas donde no las hay. Para esto se evalúan 7 problemas lineales correspondientes a la iteración número 2, etapa 20 y simulación 1. Uno de ellos es de la fase primal y 6 de la fase dual. En la Tabla Anexo 19 se muestran los resultados en un problema definido con 6 aperturas distintas y en la Tabla Anexo 20 los resultados para 3 aperturas iguales.

Problema	<i>presolve time</i> [s]	<i>solution time</i> [s]	Iteraciones
02-f1P-020-01-00.lp	0.02	0.04	378
02-f2D-020-01-01.lp	0.01	0.03	431
02-f2D-020-01-02.lp	0.01	0.03	398
02-f2D-020-01-03.lp	0.02	0.04	449
02-f2D-020-01-04.lp	0.02	0.04	442
02-f2D-020-01-05.lp	0.01	0.04	469
02-f2D-020-01-06.lp	0.01	0.04	425

TABLA ANEXO 19: TIEMPOS DE EJECUCIÓN EN CPLEX PARA CASO CON 6 APERTURAS DISTINTAS

Problema	<i>presolve time</i> [s]	<i>solution time</i> [s]	Iteraciones
02-f1P-020-01-00.lp	0	0.02	354
02-f2D-020-01-01.lp	0.01	0.02	354
02-f2D-020-01-02.lp	0.01	0.03	354
02-f2D-020-01-03.lp	0.01	0.03	354
02-f2D-020-01-04.lp	0.01	0.03	354
02-f2D-020-01-05.lp	0.01	0.03	354
02-f2D-020-01-06.lp	0.01	0.02	354

TABLA ANEXO 20: TIEMPOS DE EJECUCIÓN EN CPLEX PARA CASO CON 6 APERTURAS IGUALES

Puede verse que los problemas que incluyen aperturas replicadas tienen un menor tiempo de ejecución, además en esos casos se alcanza la convergencia en un número menor de iteraciones. De esta manera, la resolución de estos 7 problemas lineales evaluados alcanza un total de 0.36 [s] para aperturas distintas versus 0.24 [s] para aperturas replicadas, lo que significa una diferencia del 50%. Con respecto a la complejidad del problema, los problemas con aperturas replicadas parecen ser de menor complejidad; pues estos problemas son resueltos en 354 iteraciones, mientras que el menor número de iteraciones en el caso de aperturas distintas es de 378.

ANEXO D: RESULTADOS DE AGRUPACIÓN DE CLUSTERS

Como parte de los anexos se presentan los resultados de agrupamiento en 3 y 6 clusters (FIGURA ANEXO 4) para ejemplificar la selección del año representativo como el dato más cercano al centroide según la distancia euclídeana en 48 dimensiones.

También se incluyen los valores de silhouette calculados en las distintas opciones de clusters evaluadas.

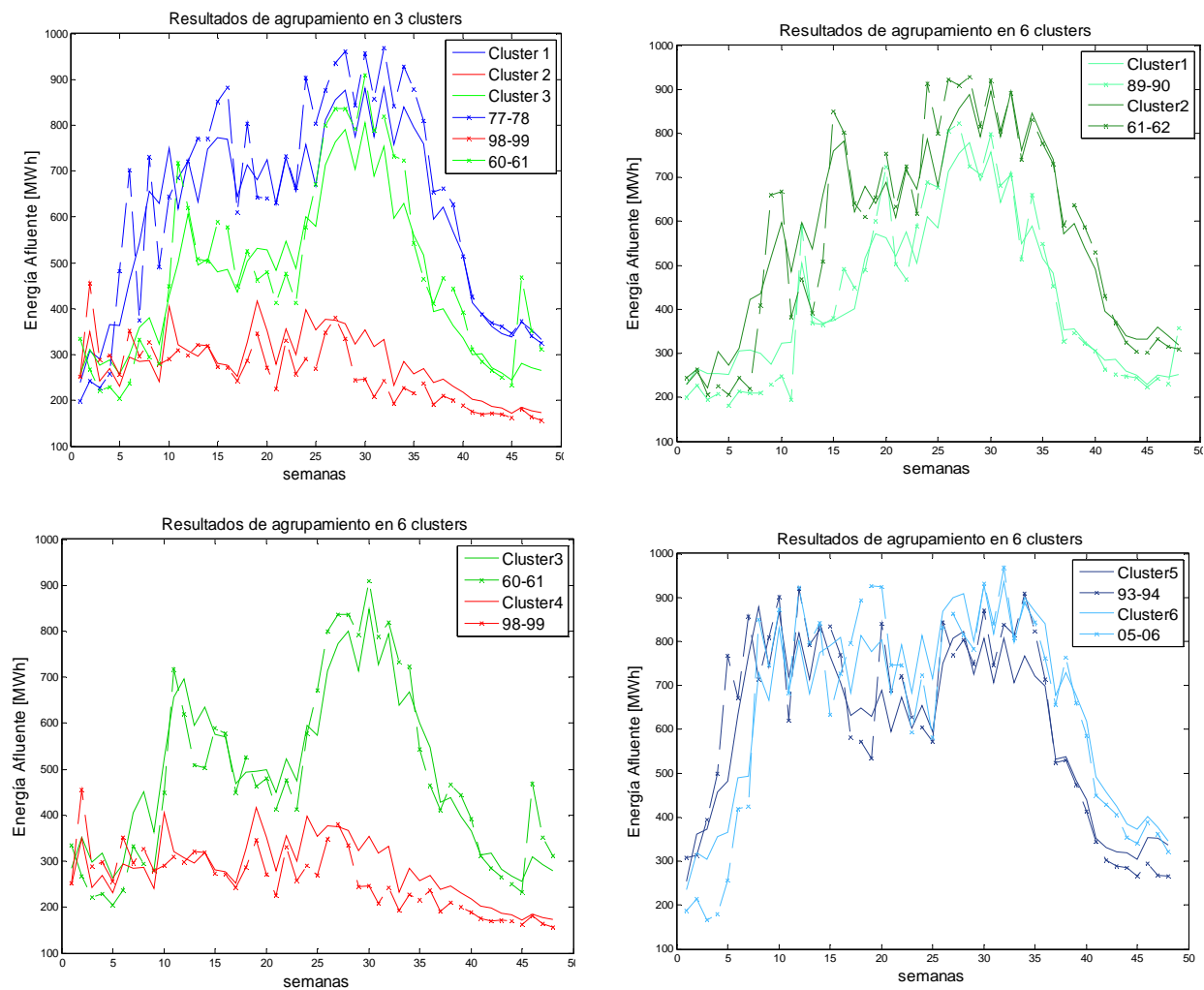


FIGURA ANEXO 4: RESULTADOS DE AGRUPAMIENTO EN 3 Y 6 CLUSTERS (3 GRÁFICOS, DOS GRUPOS POR GRÁFICO) – CENTROIDE Y AÑO HIDROLÓGICO REPRESENTATIVO

Año Hidrológico	k = 6		k = 7		k = 8		k = 9		k = 10		k = 11		k = 12	
	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)
60-61	3	0.1977	5	0.1977	3	0.2773	7	0.2773	9	0.2773	10	0.2773	2	0.2136
61-62	2	0.1768	2	0.0079	6	0.0590	4	0.0535	7	0.1587	8	0.1587	3	0.1587
62-63	1	0.1749	7	0.1749	4	0.1749	6	0.1749	4	0.1749	3	0.1749	10	0.2007
63-64	2	0.1283	2	0.2158	2	0.1667	2	0.1013	1	0.1142	1	0.1142	6	0.1142
64-65	1	0.1657	7	0.1657	4	0.1216	6	0.1216	4	0.1216	3	0.1216	10	-0.0175
65-66	6	0.2246	4	0.2148	7	0.2234	3	0.2101	6	0.1392	7	0.1392	12	0.1392
66-67	6	0.0929	4	0.0543	7	0.0567	2	0.1017	1	0.0760	1	0.0760	6	0.0760
67-68	1	0.0482	7	0.0482	4	0.0248	6	0.0248	4	0.0248	3	0.0248	2	0.1216
68-69	4	0.5131	6	0.5131	1	0.5131	8	0.5131	3	0.5131	11	0.5131	1	0.4888
69-70	2	0.0462	3	0.1464	6	0.1227	4	0.1227	7	0.0432	8	0.0432	3	0.0432
70-71	3	0.1148	5	0.1148	3	0.1702	7	0.1702	9	0.1702	10	0.1702	2	0.2434
71-72	2	0.0191	3	0.0187	6	0.0418	4	0.0418	7	0.1190	8	0.1190	3	0.0797
72-73	6	0.2198	4	0.2046	7	0.2046	3	0.2371	6	0.2409	7	0.1544	12	0.1544
73-74	3	0.1945	5	0.1258	8	0.1553	9	0.1553	10	0.1553	9	0.1553	11	0.1392
74-75	3	0.2005	5	0.1875	3	0.1704	7	0.1704	9	0.1704	10	0.1704	2	0.2227
75-76	5	-0.0605	3	-0.0205	8	0.1471	9	0.0900	10	0.0900	9	0.0373	11	0.0373
76-77	3	0.1345	5	0.1345	3	0.2244	7	0.2244	9	0.2244	10	0.2244	2	0.2517
77-78	6	0.1291	4	0.1020	7	0.0856	3	0.0840	6	0.0628	7	0.0628	12	0.0628
78-79	2	0.0754	3	0.0380	6	0.1120	4	0.0671	7	0.0985	8	0.0985	3	0.0985
79-80	2	0.0690	2	0.2293	2	0.1768	1	0.2620	8	1.0000	2	1.0000	9	1.0000
80-81	5	0.1231	1	0.1983	5	0.1983	5	0.1859	5	0.2135	6	-0.0333	7	-0.0333
81-82	5	0.1947	1	0.1844	5	0.0614	5	0.0614	5	0.0614	6	-0.0340	7	-0.0340
82-83	6	0.2199	4	0.1891	7	0.2003	3	0.1444	6	0.1583	7	0.1583	12	0.1583
83-84	3	0.2136	5	0.2136	3	0.2454	7	0.2454	9	0.2454	10	0.2454	2	-0.0020

TABLA ANEXO 21: VALORES DE SILHOUETTE PARA CADA AÑO HIDROLÓGICO (60-61 A 83-84) PARA OPCIONES DE CLUSTERS ENTRE 6 Y 12.

Año Hidrológico	k = 6		k = 7		k = 8		k = 9		k = 10		k = 11		k = 12	
	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)	Cluster	s(i)
86-87	5	0.0517	1	0.0488	5	0.0488	5	0.0329	5	0.0626	4	0.1825	5	0.1825
87-88	2	0.1215	3	0.1063	6	0.2263	4	0.2187	7	0.0782	8	0.0782	3	0.0782
88-89	1	0.0664	7	0.0664	4	0.0066	6	0.0066	4	0.0066	3	0.0066	2	0.0884
89-90	1	0.2701	7	0.2701	4	0.2201	6	0.2201	4	0.2201	3	0.2201	10	0.0813
90-91	1	0.1405	7	0.1405	4	0.1295	6	0.1295	4	0.1295	3	0.1295	10	0.0652
91-92	5	0.1080	3	0.0633	8	0.1249	9	0.1249	10	0.1144	9	0.0631	11	0.0631
92-93	5	0.0938	1	0.0926	5	0.0329	5	0.0329	5	0.0329	4	0.2308	5	0.2308
93-94	5	0.1553	1	0.2007	5	0.2007	5	0.1943	5	0.1859	4	0.2322	5	0.2322
94-95	2	0.0051	3	0.0789	6	0.0875	4	0.0875	7	0.1178	8	0.1178	3	0.1288
95-96	2	0.0305	3	0.1276	6	0.1190	4	0.1190	2	0.0779	5	0.0779	4	0.0722
96-97	4	0.4539	6	0.4539	1	0.4539	8	0.4539	3	0.4539	11	0.4539	1	0.4219
97-98	6	0.1317	4	0.0896	7	0.0600	3	0.0941	2	0.0486	5	0.0486	4	0.0486
98-99	4	0.6123	6	0.6123	1	0.6123	8	0.6123	3	0.6123	11	0.6123	1	0.5941
99-00	1	0.2023	7	0.2023	4	0.1386	6	0.1386	4	0.1386	3	0.1386	10	0.0513
00-01	2	-0.0174	3	0.0034	6	0.0346	4	0.0336	2	0.0978	5	0.0978	4	0.0978
01-02	5	-0.0268	3	0.0783	6	0.0484	4	0.0483	2	0.0562	5	0.0562	4	0.0562
02-03	6	0.1002	2	-0.0066	2	-0.0630	1	-0.0095	6	0.0348	7	0.0348	12	0.0348
03-04	3	0.2486	5	0.2349	3	0.1981	7	0.1981	9	0.1981	10	0.1981	8	0.2769
04-05	3	0.1703	5	0.1703	3	0.1934	7	0.1934	9	0.1934	10	0.1934	8	0.3341
05-06	6	0.2502	4	0.2075	7	0.2099	3	0.2289	6	0.0702	7	0.0702	12	0.0702
06-07	6	0.1806	4	0.0897	7	0.0910	3	0.1051	2	0.0649	5	0.0649	4	0.0649
07-08	1	0.2047	7	0.2047	4	0.1765	6	0.1765	4	0.1765	3	0.1765	10	0.1066

TABLA ANEXO 22: VALORES DE SILHOUETTE PARA CADA AÑO HIDROLÓGICO (86-87 A 07-08) PARA OPCIONES DE CLUSTERS ENTRE 6 Y 12.