



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

CONSIDERANDO EL CAMBIO CLIMÁTICO EN UN MODELO DE PRODUCCIÓN FORESTAL

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL
INDUSTRIAL

CAMILA BEATRIZ ZAPATA TAPIA

PROFESOR GUÍA:
ANDRÉS WEINTRAUB POHORILLE

MIEMBROS DE LA COMISIÓN:
PATRICIO CONCA KEHL
RODOLFO URRUTIA URIBE

SANTIAGO DE CHILE
2015

**RESUMEN DE LA MEMORIA PARA
OPTAR AL TÍTULO DE:** Ingeniero Civil
Industrial
POR: Camila Beatriz Zapata Tapia
FECHA: 13 de Abril de 2015
PROFESOR GUÍA: Andrés Weintraub
Pohorille

CONSIDERANDO EL CAMBIO CLIMÁTICO EN UN MODELO DE PRODUCCIÓN FORESTAL

El cambio climático puede impactar sustancialmente en el sector forestal, existen diversos estudios que demuestran que los inviernos serán más cálidos y las temporadas secas aumentarán. El objetivo de este proyecto es elaborar un método que permita incluir el factor del calentamiento global en las decisiones de producción forestal. Esto a partir de un trabajo anterior, el cual concluye que agregar incertidumbre en estos modelos aumenta el beneficio esperado a través de la inclusión de 32 escenarios climáticos, agrupados en un árbol de escenarios. El aporte de este trabajo será especificar hasta qué punto es provechoso seguir aumentando los escenarios considerados.

La complejidad del problema viene dada por la cantidad de datos a utilizar, ya que al aumentar el tamaño del árbol de escenarios, los tiempos de resolución incrementan considerablemente, esto ya que los escenarios comparten etapas, que deben tener valores comunes para las variables de decisión. Así, se obtiene un plan contingente, condicional a lo que ocurra en cada etapa, teniendo siempre la mejor respuesta.

La metodología considera generar los escenarios de clima basado en predicciones ya hechas en distintas estaciones meteorológicas de Portugal, país donde está ubicada la zona de estudio. Luego, se busca construir un modelo de programación lineal, donde la decisión corresponde a qué áreas de un bosque talar en cada periodo. Se busca maximizar los beneficios obtenidos según los precios que corresponden a cada escenario. Para llegar a una solución con una gran cantidad de escenarios, será necesario utilizar un algoritmo que permita simplificar la resolución, en este caso Progressive Hedging, cuyo principio es resolver a través de la descomposición por escenarios. Se estudian mejoras al algoritmo, además de opciones de reducción y agrupación de escenarios, para mejorar la eficiencia del mismo.

Los resultados obtenidos indican que la utilización de Progressive Hedging ofrece una eficiencia superior para instancias de gran tamaño. Esto a partir de las instancias mayores a los 300 escenarios, y además se concluye que el número óptimo de escenarios a considerar rodea los 650, en dicho punto el costo adicional para general la solución no se ve justificado, ya que la calidad de la misma no aumenta significativamente. Por otra parte, la técnica de clustering para la reducción de escenarios resultó efectiva para mejorar aún más los tiempos de resolución, considerando su correspondiente pérdida en la calidad de la solución.

Quiero agradecer primero que todo a mi familia, que siempre estuvo ahí. A mis padres que siempre me han apoyado en mis proyectos, y que desde pequeña me han motivado a perseguir mis sueños. En especial, quiero agradecer a mi hermano Fabián, gracias por tu amor a la programación y por darte el tiempo de ayudarme.

Quiero agradecer a mis amigos, Fran, Concho, Emi, Lore, Mati, Pablo, Andrés, Camilo. Gracias por preguntarme cómo iba con mi memoria y gracias por dejar de preguntarme cuando les dije que estaba estresada.

Gracias amor, por acompañarme en esta última etapa en la universidad. Me motivaste a trabajar y me diste un respiro cuando necesitaba distraerme, pero sobre todo, siempre creíste en mí.

Por último, muchísimas gracias a Cris, sin tu apoyo probablemente me hubiese demorado un par de años en terminar este trabajo. Gracias por tu ayuda y por toda tu buena onda.

Tabla de Contenido

1. Introducción.....	1
1.1 Organización del contenido	2
2. Descripción del Proyecto y Justificación	3
2.1 Objetivos.....	4
2.1.1 Objetivo General	4
2.1.2 Objetivos Específicos	4
2.2 Metodología.....	4
2.2.1 Generar escenarios futuros para la zona de estudio.....	4
2.2.2 Construcción del modelo de optimización para la producción forestal	5
2.2.3 Implementación del algoritmo de Progressive Hedging	6
2.2.4 Validación y calidad de los resultados obtenidos.....	6
2.3 Alcances	7
2.4 Resultados Esperados	7
3. Marco Conceptual.....	9
3.1 Planificación de Producción Forestal	9
3.2 Modelo de Optimización	9
3.3 Árboles y escenarios.....	10
3.3.1 Principio de No Anticipatividad.....	10
3.4 Progressive Hedging	11
3.4.1 Intuición	11
3.4.2 Pseudo-código	13
3.4.3 Caracterización.....	13
4. Problema Forestal.....	15
4.1 Antecedentes	15
4.2 Descripción del problema	15
4.3 Generación de Escenarios.....	17
4.4 Reducción del árbol de escenarios	20
4.5 Modelo de Producción Forestal	21
5. Algoritmo <i>Progressive Hedging</i>.....	24
5.1 Pseudo-código extendido.....	24
5.2 Ajustes y mejoras	25
5.2.1 Acelerar la resolución de cada subproblema.....	25
5.2.2 Linealizar el término cuadrático proximal	26

5.2.3 Ajustar el parámetro ρ de penalización	27
5.2.4 Paralelizar el algoritmo	27
5.2.5 Fijar variables.....	28
5.2.6 Variar el criterio de término	29
5.2.7 Clustering de escenarios	29
5.2.8 Bundling de escenarios	31
6. Implementación.....	33
6.1 Formulación extendida: PYOMO	33
6.2 Progressive Hedging: PYOMO - PYSP	33
6.3 Calidad de la solución	34
6.4 Soporte computacional	35
7. Resultados	36
7.1 Resultados generales	36
7.2 Comparación entre Modelo Determinístico y Estocástico.....	37
7.3 Comparación entre Formulación Extendida y Progressive Hedging.....	38
7.4 Caso clustering de escenarios	40
7.4.1 Conglomerados de K medias.....	40
7.4.2 Clusters de tamaño uniforme.....	41
7.5 Caso bundling de escenarios.....	42
7.6 Factibilidad y Calidad de la Solución.....	43
7.6.1 Factibilidad para el caso de Clustering	44
8. Conclusiones.....	48
8.1 Trabajo futuro	49
9. Glosario	51
10. Bibliografía.....	52
11. Anexos	53
Anexo A. Teorema de Redistribución de probabilidades.....	53
Anexo B. Extracto de los datos entregados por el programa ForEAdapt.....	53
Anexo C. Resultados para distintos niveles de demanda para todas las instancias.....	55

Índice de Figuras

Figura 3-1 Escenarios por separado y escenarios compuestos en un árbol	10
Figura 4-1 Bosque Chouto Parreira y su ubicación en Portugal.....	16
Figura 4-2 Ejemplo para selección de climas extremos	17
Figura 4-3 Niveles mínimos y máximos de precipitación por periodo	18
Figura 4-4 Ejemplo de posibles combinaciones entre nodos.....	19
Figura 4-5 Árbol para una muestra de 100 escenarios	19
Figura 4-6 Representación gráfica de una decisión para instancia de prueba de 55 escenarios, en el periodo 10	23
Figura 5-1 Diagrama de linealización por tramos.....	26
Figura 5-2 Ejemplo de clusters de tamaño uniforme	31
Figura 5-3 Ejemplo de bundling de escenarios.....	32
Figura 7-1 Resultados para distintos niveles de demanda para instancias de 1, 10, 200, 500, 800 y 1240 escenarios	37
Figura 7-2 Comparación entre Modelo Determinístico y Estocástico	38
Figura 7-3 Comparación entre Formulación Extendida y Progressive Hedging	39
Figura 7-4 Gráfico para los resultados de la comparación	39
Figura 7-5 Resultados de Clustering con conglomerados de K medias para instancias pequeñas	40
Figura 7-6 Resultados de Clustering con conglomerados de K medias para instancias masivas	41
Figura 7-7 Resultados de clustering de tamaño uniforme	42
Figura 7-8 Nivel de infactibilidad por instancia	43
Figura 7-9 Gráfico para el factor de infactibilidad	44
Figura 7-10 Factibilidad para instancias reducidas por clustering.....	45
Figura 7-11 Gráfico de infactibilidad para caso Clustering con K medias	46
Figura 7-12 Gráfico de infactibilidad para caso Clustering uniforme	46
Figura 11-1 Extracto de los datos para una estación meteorológica.....	53
Figura 11-2 Resultados para distintos niveles de demanda para todas las instancias	55

Capítulo 1

Introducción

El cambio climático puede impactar sustancialmente en el sector forestal. Existen diversos estudios que indican que los inviernos serán más cálidos y que tanto el largo como la frecuencia de las temporadas secas aumentarán.

En particular, este trabajo se enfocará en bosques de eucaliptos de Portugal. Esta especie representa el 26% de todos los bosques de dicho país según el último Inventario Forestal Portugués [1], siendo la principal materia prima en la industria de la celulosa y del papel.

La producción forestal de los eucaliptos se caracteriza por considerar una rotación cercana a los 12 años, esto implicaría que ante cambios climáticos en ese periodo, la planificación anterior llevaría a ingresos diferentes a los esperados.

Hasta el momento, se han desarrollado modelos de producción forestal, pero principalmente de carácter empírico, donde se planifica considerando un solo escenario de clima, que corresponde al caso de las condiciones históricas de la zona. También se ha considerado encontrar la producción óptima para un escenario particular, que pueda considerar el cambio climático, pero es sabido que el clima es estocástico.

El desarrollo del proyecto que se está presentando está enmarcado en el programa ForEAdapt (Knowledge Exchange between Europe and America on forest growth models and optimization for adaptive forestry) de la Comunidad Europea [2], que apoya la colaboración entre estudiantes de distintos países para generar modelos de producción forestal.

Esta institución apoya la realización del proyecto entregando los datos necesarios para desarrollar los modelos, según los requerimientos que se presenten. Estos datos corresponden a predicciones de crecimiento de un bosque particular, simulados ya por expertos climatólogos que participan del programa.

En la línea de estudio del presente trabajo, existen avances en el tema, que buscan corroborar la tesis de que al agregar escenarios de clima los beneficios esperados aumentan. En particular, este trabajo de título buscará continuar con la investigación realizada por Jordi García y Andrés Weintraub, investigador parte del programa ForEAdapt, que trabajó con una muestra de 32 escenarios de clima para la zona de estudio [3]. En ese primer intento, se concluye que efectivamente el beneficio esperado aumenta al considerar el factor estocástico.

El aporte de este trabajo de título será ampliar el rango de estudio de los escenarios climáticos, con la complejidad que eso conlleva, en un horizonte de tiempo similar a los trabajos ya realizados, y así generar conocimiento que pueda ser comparable.

1.1 Organización del contenido

La memoria se desarrolla de la siguiente manera: en el capítulo 2 se describe el proyecto, con todos sus componentes, objetivos, metodología y resultados esperados. En el capítulo 3 se muestran los conceptos necesarios para familiarizarse con la optimización estocástica y una introducción al algoritmo que se utiliza para resolver el problema. En el capítulo 4 se detalla el problema forestal, los antecedentes, la generación de escenarios y la formulación del problema. Luego, en el capítulo 5 se describe el algoritmo de solución para problemas de este tipo y los ajustes que deben hacerse, posteriormente en el capítulo 6 se detalla cómo se implementarán las metodologías planteadas. Así, en el capítulo 7 se muestran los resultados generales, la comparación del modelo extendido con el algoritmo planteado, y también un análisis de calidad de los resultados. Finalmente, en el capítulo 8 se desarrollan conclusiones y se plantean desafíos futuros.

Capítulo 2

Descripción del Proyecto y Justificación

El proyecto que se va a desarrollar consiste en elaborar un modelo de producción forestal, que considere el efecto del calentamiento global a través de la inclusión de variables estocásticas. El carácter estocástico estará dado por la consideración de escenarios climáticos. A diferencia de lo ya realizado en [3] considerando 32 escenarios, se buscará incluir un número mucho mayor de escenarios, y así representar de una forma más exacta la alta gama de posibles climas. Uno de los objetivos es concluir sobre los beneficios de considerar tantos escenarios, comparando las ganancias en cuanto a precisión del modelo en contraposición a un alto costo de calcular las variables de la planificación.

Se toma en cuenta una planificación de mediano plazo para un bosque portugués, ZIF Chouto Parreira. Ubicado en la zona central de Portugal, este bosque está dominado por la especie eucalipto, área que debería ser completamente cosechada en un periodo de 15 años.

Las características de los escenarios a estudiar corresponden a datos entregados por el programa ForEAdapt. Cada escenario está definido por el valor de los siguientes atributos: temperatura media, temperatura máxima y mínima, precipitación total, radiación solar, días de lluvia, días de helada, presión atmosférica, evaporación y humedad. Datos que predicen lo que ocurriría entre los años 2014 a 2030, a nivel mensual. Los escenarios corresponden a predicciones en estaciones meteorológicas, en distintas ubicaciones geográficas de Portugal. A partir de estas predicciones se genera un árbol de escenarios, que abarca condiciones similares a zonas más australes y a zonas al norte del bosque en estudio, de forma de considerar las posibilidades de cambio climático que experimentará el país.

Luego, a partir de estos escenarios, se hace una simulación de lo que ocurriría en el periodo de estudio, indicando la productividad de las zonas de tala y el precio de venta en cada caso. Así se puede obtener el ingreso para cada área, según la proporción de la misma que se va a producir en cada periodo. La decisión que deben tomar las empresas forestales considera qué zonas de sus bosques talar en cada periodo y qué proporción de las mismas. Existen consideraciones que se deben tomar, como la demanda que tienen, la productividad de sus plantaciones y además, restricciones intrínsecas de la tala de árboles.

Considerando las metodologías a utilizar, tomando en cuenta muchos escenarios en varias etapas temporales, el problema no puede ser resuelto directamente usando cualquier Solver comercial. Esto principalmente por las restricciones de no anticipatividad, y el tamaño de las instancias que se buscará resolver. En el problema tratado en esta memoria, se conformará todo el universo de climas futuros en un árbol de escenarios. Así, se descartan metodologías como Relajación Lagrangeana o Programación Dinámica Estocástica, en las cuales trabajar con muchos escenarios y múltiples etapas se hace demasiado complejo. Una hipótesis que será probada en este trabajo será la de utilizar el algoritmo de Progressive Hedging, y que éste será capaz de entregar resultados óptimos, esta técnica ha demostrado ser conveniente en otros trabajos similares y por eso se opta por esta metodología.

Finalmente, cabe recalcar la relevancia de este trabajo. Existen modelos que consideran variables estocásticas en forma de escenarios de clima, a la hora de tomar decisiones de producción forestal. El aporte en este caso es estudiar el beneficio y la factibilidad de considerar un rango mucho más amplio de escenarios, que se acerca al ideal que es un continuo. Además, el valor que agrega no solo recae en los resultados que se puedan obtener, sino también las distintas aplicaciones que eventualmente pueda conllevar. Se obtienen escenarios para zonas de Portugal, pero igualmente es interesante en el sentido que el modelo generado puede en algún momento ser utilizado en otras zonas del mundo, por ejemplo, con un clima similar.

2.1 Objetivos

A continuación se presenta el objetivo general del trabajo de título, y el desglose del mismo en 6 objetivos específicos.

2.1.1 Objetivo General

Elaborar un método que permita incluir el factor del calentamiento global en las decisiones de producción forestal, a través del análisis de una gran cantidad de escenarios climáticos y sus probabilidades de ocurrencia, así se tendrá una planificación que responderá ante las incertidumbres de clima.

2.1.2 Objetivos Específicos

1. Generar escenarios futuros de clima para la zona de estudio como una combinación de factores que afectan la productividad de los bosques, a partir de predicciones planteadas por climatólogos.
2. Generar un modelo de optimización, y lograr su correcta respuesta frente a la gran cantidad de escenarios generados.
3. Implementar correctamente el algoritmo de Progressive Hedging, de forma de encontrar la configuración del mismo que permitirá trabajar con gran cantidad de escenarios en un tiempo razonable.
4. Evaluar el modelo desarrollado y verificar su factibilidad, debe entregar soluciones aplicables en todos los escenarios simulados.
5. Concluir sobre la cantidad de escenarios óptima a utilizar, contrastando el utilizar pocos escenarios para mejorar en el tiempo y utilizar muchos escenarios para mejorar la calidad de la solución.
6. Cuantificar el beneficio que implica usar este tipo de metodologías, versus un modelo empírico como los que se utilizan en la actualidad.

2.2 Metodología

La metodología determinada para realizar este proyecto se divide en 4 actividades principales, que se describen a continuación, especificando cómo se espera realizar cada una de ellas.

2.2.1 Generar escenarios futuros para la zona de estudio

Esta metodología entregará los escenarios que luego servirán para simular la producción futura de la zona a estudiar en Portugal. Un escenario corresponde al clima para un año, que corresponde a un periodo del horizonte de evaluación.

Los niveles de productividad, para cada área de tala, en cada periodo del horizonte de evaluación, en conjunto con el precio de venta, servirán como el principal insumo para tomar las decisiones de producción. El objetivo es que los escenarios a utilizar representen de buena forma un extenso rango de posibilidades de clima, y ver si esto mejora los ingresos totales percibidos.

Para realizar esta actividad, se cuenta con predicciones de clima para 15 años de estaciones meteorológicas ubicadas en distintas zonas de Portugal, basados en EMSEMBLES Project [9] que proveen de bases de datos de clima desarrollados por el Centro Hadley. Los escenarios se construirán haciendo una combinación de estos climas, de forma que correspondan a condiciones factibles de ocurrir y con una combinación posible de los factores que definen un escenario.

Por otra parte, se va a considerar solamente el factor del nivel de precipitación de los tres principales meses de crecimiento de los eucaliptos, correspondiente a los meses de Marzo, Abril y Mayo. Es decir, un año de un escenario estará caracterizado por la suma de las precipitaciones en esos tres meses.

El resultado esperado, es un árbol de escenarios. Para cada periodo, existirán nodos intermedios entre el nivel máximo y el mínimo de precipitación. Para que los nodos resultantes correspondan a escenarios factibles de ocurrir, conociendo el valor de la precipitación en cada nodo, se buscarán climas con dicho nivel de lluvia y se utilizará el valor de los otros factores para terminar de definir ese escenario.

Se considerará que dado que los saltos de climas no deben ser demasiado marcados en periodos cercanos, cada nodo del árbol solo se combinará con los nodos adyacentes (hacia arriba y hacia abajo).

Una vez obtenido el árbol, será necesario reducir la dimensión del problema, y así contar con distintas instancias de prueba y poder analizar respecto a las soluciones obtenidas para distintos números de escenarios. Para esto, se utilizarán algoritmos de reducción de escenarios que permitan mantener la distribución de probabilidad del árbol original.

Finalmente, a partir de los escenarios generados, se envían los datos a un simulador de productividad. Este simulador es entregado por ForEAdapt, que toma el nivel de precipitación (además de las otras condiciones climáticas que afectan el crecimiento de los bosques) y entrega los valores de precio y productividad para cada caso, en cada área de tala y en cada periodo.

2.2.2 Construcción del modelo de optimización para la producción forestal

Para la construcción del modelo, se estudiará en primer lugar el modelamiento ya realizado en otros trabajos del programa ForEAdapt. El objetivo es considerar todos los factores clave en la producción forestal, en relación a satisfacer la demanda, cumplir con ciertas condiciones medioambientales, entre otras. De acuerdo a lo investigado se definirán los elementos que incluye el modelo:

- Condición de optimización: se define si el modelo corresponde a maximización o minimización.
- Función objetivo: con qué criterio se tomarán las decisiones, generalmente dada por costos, beneficios, ingresos, entre otros.
- Variables de decisión: representan las decisiones que se desea apoyar. Corresponderían por ejemplo a qué áreas talar en cada periodo.

- Restricciones: son las condiciones que deben cumplir las decisiones, las limitaciones inherentes al problema.
- Parámetros del problema: son los datos que están dados, es decir, no son controlables por el tomador de decisión.

Finalmente, una vez construido el modelo, se debe encontrar la aplicación que lo resuelva para obtener el valor de las variables de decisión, y también el beneficio percibido. En este caso es de vital importancia considerar el tamaño del modelo, en relación al número de variables, la cantidad de parámetros, el número de restricciones, entre otros.

En relación al software, se utilizará PYSP del repositorio Coopr de los Laboratorios Nacionales Sandia [4]. Paquete en lenguaje Python de optimización, que cuenta con facilidades para expresar modelos abstractos, es decir, que maneja de forma independiente las instancias. Adicionalmente, cuenta con la implementación directa de Progressive Hedging, que ya se destacó como el algoritmo a utilizar para resolver con el gran número de escenarios.

2.2.3 Implementación del algoritmo de Progressive Hedging

Una hipótesis que debe ser probada es que la utilización del algoritmo Progressive Hedging permitirá trabajar con un número mucho mayor de escenarios, en comparación a trabajar con la versión extendida del problema.

En primer lugar, se debe concluir en qué punto el soporte computacional disponible permitirá trabajar la formulación extendida del problema forestal, es decir, definir el número de escenarios con que el problema extendido “se cae” y por lo tanto será necesario buscar alternativas de resolución.

Luego, y de forma paralela a la búsqueda de ese número máximo de escenarios para el problema extendido, se resolverá para las distintas instancias el mismo problema utilizando Progressive Hedging, de forma de poder comparar el rendimiento del algoritmo, y a su vez encontrar la configuración correcta de sus parámetros (estos serán especificados en las próximas secciones). Se espera que el tiempo que le tomará al algoritmo solucionar el problema pueda variar en gran medida al implementar distintas técnicas para acelerar la convergencia.

Finalmente, el objetivo será encontrar hasta qué punto el algoritmo permitirá resolver instancias de mayor tamaño y por lo tanto, si se desea seguir aumentando el número de escenarios, en qué punto buscar nuevas alternativas de resolución. Es importante en este punto también, analizar y comparar con resultados de otros trabajos por qué es importante agregar escenarios y a qué nivel es significativo seguir aumentando. Se deberá concluir cuál sería el número óptimo de escenarios de forma de cubrirse frente a la incertidumbre y a la vez ser capaz de resolver el problema en un tiempo razonable y con los recursos computacionales disponibles.

2.2.4 Validación y calidad de los resultados obtenidos

Al trabajar con una gran cantidad de escenarios, utilizando muchos recursos computacionales y de tiempo, es importante concluir sobre la utilidad de los modelos obtenidos. Ya se ha concluido en trabajos anteriores sobre las mejoras que se obtienen al comparar modelos determinísticos frente a aquellos que incluyen factores estocásticos. En este trabajo, se centrará en comprobar la utilidad de aumentar la cantidad de escenarios, que representan el factor

estocástico del modelo, y a la vez concluir sobre la cantidad óptima de escenarios a utilizar bajo estas condiciones.

La metodología a utilizar en este punto toma en primer lugar la instancia de mayor tamaño con la que se está trabajando. En este caso, corresponde a una instancia de 1240 escenarios. Luego, al tener las soluciones para las instancias de menor tamaño, se desea corroborar qué parte de los 1240 escenarios originales son infactibles bajo las soluciones de las instancias menores, es decir, que no puedan cumplir con las restricciones de producción y demanda, dada la productividad de cada uno de esos escenarios.

Finalmente, la medida de calidad de cada solución será el porcentaje de escenarios de la instancia de 1240 que son infactibles una vez obtenida la solución para cada una de las instancias menores. Con esos datos obtenidos, se espera poder concluir la cantidad de escenarios óptima.

2.3 Alcances

El proyecto que se está presentando abarca desde la generación de escenarios de clima, basados en predicciones hechas por climatólogos relacionados al programa ForEAdapt. Posteriormente, a partir de estos escenarios, se realiza una simulación del crecimiento del bosque en la zona de estudio, junto con el precio de venta de la potencial producción en ese periodo. Esta simulación es llevada a cabo también por expertos del programa, que tomando el conjunto de características de cada escenario, concluye cuál sería el resultado en cuanto a productividad.

Además, cabe considerar que la información entregada para el desarrollo de este proyecto corresponde a una única zona de bosques, en Portugal. Todo esto con la productividad correspondiente a esas latitudes y con precios de ventas para los mercados locales de ese país. Para ser adaptado a la realidad de otros países, es necesario considerar que se requieren esos datos, lo que no será trabajado en esta ocasión.

Luego, debe considerarse el análisis de los escenarios simulados, evaluar la cantidad de escenarios que resulta significativa a la hora de generar el modelo de producción forestal y que además resulte factible de encontrar solución en tiempos razonables y con los recursos computacionales disponibles.

Finalmente, se elabora un método que permita, usando la información proporcionada, generar un plan de decisiones para la producción forestal, en un horizonte de 15 años. El horizonte de evaluación corresponde a un requerimiento del programa en el que se enmarca este trabajo. El objetivo primordial del programa es el intercambio de conocimientos entre investigadores europeos y americanos, de forma que se exige un horizonte de planificación en común, así puede realizarse una comparación posterior, y obtener más información de los modelos generados.

2.4 Resultados Esperados

Considerando los objetivos planteados y la metodología a utilizar, se espera en primer lugar, generar un árbol de escenarios que cubra el más amplio espectro de climas esperable. Esto quiere decir que una vez generado el modelo de producción, éste deberá contar con planes contingentes para prácticamente cualquier clima que pueda presentarse en el futuro, en un horizonte de 15 años.

En relación a la resolución del problema, se espera definir en qué punto es necesario encontrar alternativas para encontrar el óptimo de la formulación extendida, en relación a estudios anteriores, se prevé que el problema extendido no pueda ser resuelto para instancias mayores a los 500 escenarios. Por otro lado, se buscará una configuración óptima del algoritmo Progressive Hedging para las instancias que no puedan ser resueltas sin usar el algoritmo, así mejorar sustancialmente los tiempos de resolución del problema. En relación a eso, se espera que se pueda encontrar una solución con mejoras de hasta un 80% en el tiempo de convergencia del algoritmo para una configuración óptima.

Además, se espera concluir para qué cantidad de escenarios el algoritmo PH no sigue funcionando, esperándose que más allá de los 1000 escenarios no llegue a la convergencia exigida. También, se buscará qué cantidad de escenarios resulta óptima, esto es una relación entre la calidad de la solución y los tiempos que toma encontrar dicha solución. Respecto a este último punto no hay buenas referencias de cómo llegar a ese valor, y por lo tanto será interesante concluir al respecto.

Finalmente, resultará vital concluir sobre la factibilidad de las soluciones encontradas. Se espera que en las instancias de mayor tamaño, superiores a los 100 escenarios, se esté cubierto frente a la incertidumbre y por lo tanto para cualquier clima futuro existirá una solución aplicable al caso.

Capítulo 3

Marco Conceptual

3.1 Planificación de Producción Forestal

Para abordar el problema, se hace una investigación bibliográfica del estado del arte actual en relación a los modelos de planificación de la producción forestal. El modelo que se utiliza en la actualidad, corresponde a la versión determinística del problema. Para hacer la planificación de la tala, se considera que durante el próximo horizonte de evaluación el clima fluctuará en torno a un escenario de clima similar a los datos históricos de la zona.

Al existir parámetros estocásticos en el problema, como el crecimiento de los árboles y el precio de venta, la formulación determinística asume el valor de dichas variables como el valor promedio y luego optimiza en base a eso. Finalmente, este método provee de una solución óptima para el escenario promedio de ocurrencia.

Pero existen diversas fuentes de incertidumbre en la industria forestal, y considerando el cambio climático que está viviendo el planeta, ignorar estos efectos podría generar graves consecuencias en la rentabilidad de las forestales. La fuente de incertidumbre que trata de ser incluida por el actual trabajo, corresponde a las variaciones en el crecimiento y por consiguiente, en las variaciones del rendimiento de los bosques. En el pasado, como ya se expresó, se han modelado casos donde este efecto se ataca explícitamente, pero el rango de escenarios de clima no es amplio y por lo tanto se especula que no esté rescatando el efecto del calentamiento global.

3.2 Modelo de Optimización

De acuerdo al marco metodológico definido en “Optimización y Modelos para la Gestión” de Samuel Varas [5], aplicado a la optimización de la producción forestal, se describen las principales etapas del modelamiento.

- 1. Definición del sistema:** Corresponde a identificar el ámbito del sistema bajo estudio, establecer el objetivo de éste, identificando las alternativas de decisión.
- 2. Construcción del modelo:** La investigación operativa se caracteriza por la utilización de modelos matemáticos en la formulación de problemas. Un modelo matemático es una representación idealizada de una situación concreta con un objetivo determinado.
- 3. Resolución del modelo:** Determinar la solución de un modelo matemático implicar establecer los valores de las variables de decisión de manera que se cumplan las limitaciones o restricciones consideradas, y que optimice la medida de efectividad asociada.
- 4. Validación del modelo:** Se analiza la capacidad del modelo para predecir razonablemente el desempeño de sistema ante diversas alternativas de decisión. Luego es necesario verificar si la solución es razonable, realizar pruebas de consistencia, validación de

supuestos, entre otros. Y por último si el modelo y su solución no se consideran satisfactorios, se deben realizar las correcciones necesarias.

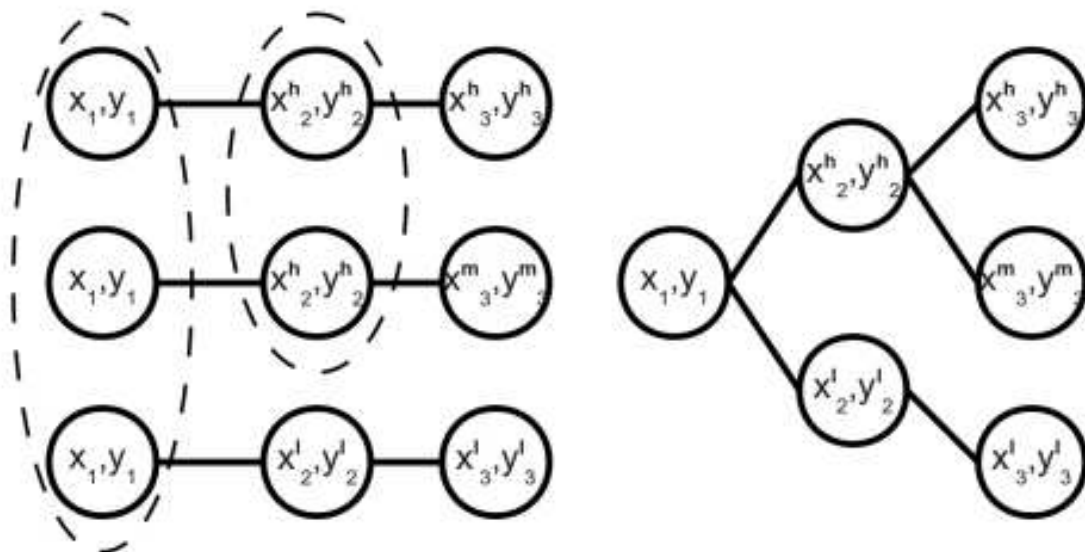
3.3 Árboles y escenarios

En este trabajo, la forma de representar la incertidumbre será a través de escenarios y la forma funcional utilizada para medir el riesgo es la esperanza. La forma en que se expresarán las decisiones será un plan de acción, condicional a lo que suceda en el futuro, a medida que avanzan las etapas. Este plan de acción estará reflejado en un árbol de decisión frente a un futuro incierto.

Bajo el concepto de optimización estocástica, un escenario es el conjunto de los valores numéricos determinados para todas las fuentes de incertidumbre, en este caso, el nivel de productividad de la zona de estudio bajo distintos climas y el precio de venta del producto extraído.

Luego, se calcula o asigna una probabilidad de ocurrencia a cada uno de los escenarios. Para formar el árbol de escenarios, se considera que todos tienen al menos una etapa en común, correspondiente a la primera: el nodo de origen. Este sería una representación de las condiciones del presente, con valores ya conocidos. Se aprecia en la Figura 3-1 una composición de un árbol de escenarios, con el nodo origen en común.

Figura 3-1 Escenarios por separado y escenarios compuestos en un árbol



3.3.1 Principio de No Anticipatividad

El principio de No Anticipatividad, como fue formulado por Wets y Rockefellar en [6], se puede comprender fácilmente de la siguiente manera: “Si dos escenarios son indistinguibles hasta cierta etapa, entonces las decisiones en estas etapas, hasta esa etapa, deben ser las mismas”. Es justamente este principio el que genera la mayor parte de la complejidad para resolver el problema planteado en este trabajo. La intuición es que si se trataran de escenarios independientes entre sí, el problema puede ser fácilmente resuelto como una combinación de

problemas más simples, pero al cumplirse este principio, hay restricciones que ligan a estos escenarios.

Formulaciones extendidas

Una formulación extendida hace referencia a que se resuelven de manera simultánea todos los escenarios. Para estos problemas, existen dos familias de formulaciones, caracterizadas a continuación:

- Explícita: variables por escenario con restricciones para la no anticipatividad.
- Compacta: variables por nodo o paquete de información, que implementan la no anticipatividad de manera implícita.

En el primer caso, tanto la cantidad de variables como de restricciones es mucho mayor, ya que para cada escenario se escogen los valores de las variables que luego deben cumplir con la no anticipatividad, en cambio en la versión compacta, las variables implementan estas restricciones implícitamente, al tomar solo un valor por nodo.

Esto lleva a que la formulación compacta tenga un mejor desempeño, y es la que se utilizará en esta memoria.

3.4 Progressive Hedging

El algoritmo de Progressive Hedging es un algoritmo de descomposición para problemas estocásticos con escenarios. Su utilidad recae en que el problema en su forma extendida, es decir con las restricciones de No Anticipatividad, es muy difícil de resolver por su gran tamaño, por lo que se divide en problemas más pequeños.

PH presenta como propiedad el ser exacto para problema lineales convexos, es decir, converge al óptimo global. Sin embargo, para problemas enteros esta condición no se garantiza, por lo que no se puede asegurar la optimalidad de la solución factible entregada.

3.4.1 Intuición

La intuición que subyace en esta descomposición consiste en la posibilidad de resolver el problema estocástico original a partir de la resolución independiente de cada uno de los escenarios modelados y luego calcular la esperanza de la solución global obtenida. Así, se resolverán $|S|$ (número de escenarios) problemas sencillos y determinísticos, lo que permite un mayor rendimiento en términos computacionales. No obstante, intercambiar la minimización con el cálculo de la esperanza no es posible debido a la presencia de las restricciones de no anticipatividad en el modelo, dado que éstas acoplan las decisiones a tomar entre los diversos escenarios.

Sin embargo, al observar la estructura de las soluciones factibles obtenidas del problema, a partir de las restricciones de no anticipatividad se tiene que:

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

Esto equivale a señalar que es posible reformular la restricción de no anticipatividad notando que todas las variables que se encuentran en cada uno de los nodos del árbol de escenarios son iguales, y por lo tanto, iguales a la esperanza de cada uno de los conjuntos de información.

Dado lo anterior, es posible desacoplar la dependencia inherente entre los diversos escenarios utilizando la técnica de relajación lagrangeana sobre este grupo de restricciones. Se supondrán conocidos los multiplicadores asociados (los penalizadores \widetilde{w}_s que penalizan la función objetivo en la medida que la solución se desvíe de la restricción de no anticipatividad), lo que implica que la función objetivo se convierte en:

$$\min_{x_s \in Q_s, \forall s \in S} \sum_{s \in S} p_s f_s(x_s) + \sum_{s \in S} \widetilde{w}_s (x_s - E(x_s))$$

Realizando un cambio de notación y reagrupando términos ($\widetilde{w}_s = p_s w_s$ y $E(w) = 0$), se obtiene:

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

Así es posible realizar el intercambio entre la minimización y el cálculo de la esperanza, lo que permite descomponer el problema original en una serie de problemas más sencillos y pequeños.

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

En base a lo anteriormente realizado, el paso final consiste en realizar el cálculo iterativo de los multiplicadores w presentes en la función objetivo. Sin embargo, su convergencia bajo esta formulación no está garantizada, por lo que para evitar esta dificultad (que podría llevar a convertir los sub-problemas en no acotados), se agrega un término proximal cuadrático que obliga a las variables X a converger en la dirección correcta. Por lo tanto, se procede a utilizar la relajación lagrangeana aumentada, cuya característica principal consiste en el uso de un término cuadrático para llevar a cabo la penalización, la cual corresponde a la diferencia que presenta cada una de las posibles soluciones obtenidas en relación al promedio de las soluciones calculadas en la iteración previa, obteniéndose así finalmente una función objetivo como la que sigue:

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

Importante es notar que en caso de que no se incluya este término cuadrático se podría producir que al realizar la minimización de la función objetivo, alguno de los multiplicadores w genere que alguna de las variables sea no acotada y así el problema no lograría converger.

3.4.2 Pseudo-código

Paso 1: Resolver cada escenario.

$$\min_{x_s \in Q_s} f_s(x_s)$$

Paso 2: Calcular la solución global en cada nodo.

$$\bar{x} = \sum_{N_t: s \in N_t} p_s x_s$$

Paso 3: Si las soluciones se parecen lo suficiente, parar.

$$\|x - \bar{x}\| < \varepsilon$$

Paso 4: Actualizar la penalización.

$$w = \rho(x - \bar{x}) + w_{-1}$$

Paso 5: Resolver cada escenario penalizado.

$$\min_{x_s \in Q_s} f_s(x_s) + x_s w_s + \frac{\rho}{2} \|x_s - \bar{x}\|^2$$

Paso 6: Ir al paso 2.

En términos sencillos, el algoritmo busca resolver de forma iterativa cada uno de los escenarios, penalizando en cada iteración con respecto a la diferencia que presenta la solución obtenida en relación al promedio o esperanza de todos los escenarios, es decir, se penaliza con respecto a la distancia que se tiene al óptimo global. Luego, la convergencia del algoritmo se medirá como la distancia euclídeana entre la solución de cada uno de los escenarios y la solución global. El algoritmo finaliza cuando todas las soluciones obtenidas han alcanzado un criterio de cercanía suficiente (iguales o muy cercanas según se defina el parámetro ε) en relación a la solución global.

3.4.3 Caracterización

Las principales características del algoritmo Progressive Hedging son:

- Es un algoritmo de descomposición horizontal (es decir, por escenarios, la descomposición vertical es por etapas). En general, estos problemas han sido tratados con descomposición vertical, por lo que ésta sería una innovación.
- No está limitado por el número de escenarios considerados, como sí lo está la formulación extendida que crece mucho más rápidamente en dificultad al incrementar los escenarios.

- Relaja la restricción de no anticipatividad poniéndola en la función objetivo, más un término proximal cuadrático.
- Posee una demostración basada en Relajación Lagrangeana Aumentada, lo que la hace exacta en el caso convexo.
- Es una heurística para el caso que involucre variables enteras.
- Presenta una estructura fácilmente paralelizable, pues basta resolver cada subproblema en distintos núcleos de forma concurrente. Así, es posible utilizar programación distributiva directamente en un clúster computacional.
- Su desempeño está fuertemente ligado a lo rápido que se pueda resolver cada uno de los subproblemas (escenarios), dado que el resto consiste básicamente en calcular promedios y sumas lineales.

Capítulo 4

Problema Forestal

4.1 Antecedentes

Un ejemplo de la introducción de la estocasticidad en los modelos de planificación lo presenta Quinteros en [10], donde se incluye la variabilidad en el mercado a través de 16 escenarios, donde varía tanto el precio como la demanda. Así, cada escenario es una representación de todos los valores de los parámetros estocásticos para cada periodo.

El actual trabajo se basa en el problema presentado por García en [3], que considera la incertidumbre en el crecimiento futuro de los bosques y la madera producida como resultado del cambio climático, esto basado en simulaciones para el volumen futuro que serán talados y al precio que deben ser vendidos. Al igual que el presente trabajo, la variabilidad inherente en el crecimiento de los árboles no es considerada.

La incertidumbre está representada por escenarios, que incluyen valores para las variables estocásticas para cada periodo de tiempo. En cuanto a la probabilidad de ocurrencia, se considera que los escenarios son equiprobables. Se escogen 32 escenarios, basados en ENSEMBLES Project [9] que proveen de bases de datos de clima desarrollados por el Centro Hadley que son considerados como los más apropiados para las condiciones en Portugal.

En el actual trabajo, se busca generar un árbol de escenarios que cubra un espectro mayor, tomando como input una mayor cantidad de bases de datos y que además la distribución de los escenarios sea más homogéneos, de forma que se pueda esperar que ante cualquier clima que se presenten, exista algún escenario lo suficientemente cercano.

Las decisiones involucran qué unidades talar en cada periodo, problema largamente estudiado, pero que no había sido estudiado bajo las incertidumbres del cambio climático. Estudios anteriores habían considerado un modelo determinista, que en general asumen que las condiciones futuras serán similares al año anterior.

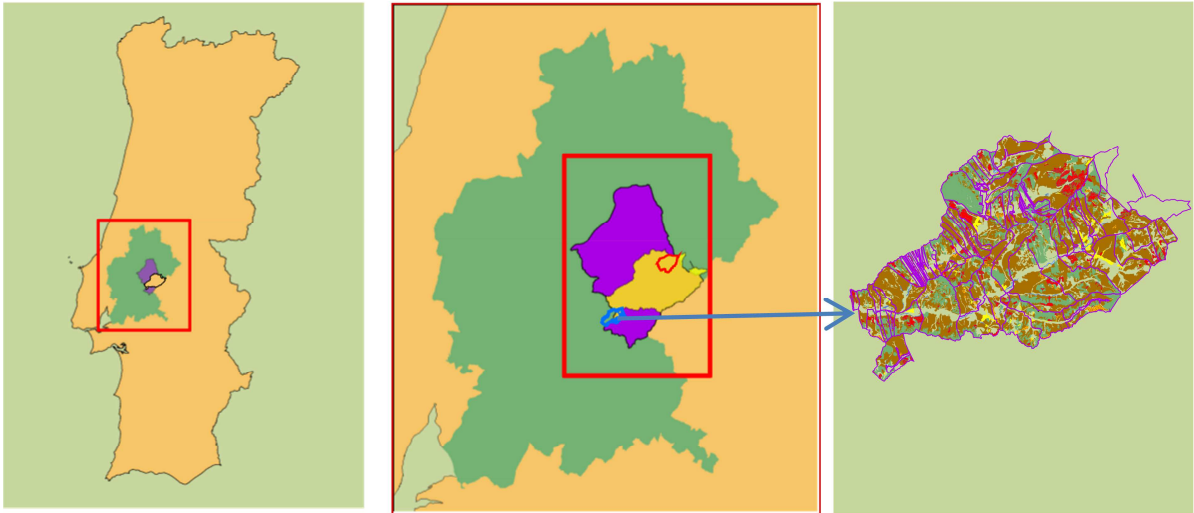
Otro factor que se debe considerar, es la discretización del bosque para tomar las decisiones. García agrupa en 21 unidades de cosecha, mientras que el actual trabajo busca decisiones más precisas, de forma que el bosque considerará las 1000 zonas en las que se divide actualmente el área de estudio.

4.2 Descripción del problema

El problema en estudio busca conocer las mejores decisiones relacionadas a la parte que debe ser explotada de cada unidad de tala de un bosque en Portugal, tomando en cuenta la variabilidad del crecimiento de cada zona debido a cambios en el clima y por lo tanto variaciones en la productividad. El objetivo consiste en maximizar el beneficio esperado, tomando en cuenta las probabilidades de ocurrencia de cada escenario climático.

Se toma en cuenta una planificación de mediano plazo para un bosque portugués, ZIF Chouto Parreira, dominado por la especie eucalipto. En la Figura 4-1 se muestra una figura de del área de estudio y su ubicación dentro del país.

Figura 4-1 Bosque Chouto Parreira y su ubicación en Portugal



Considerando el horizonte de tiempo, se asume que el bosque es apto para cosecha en su totalidad y que debe ser completamente talado en un periodo de 15 años. Así, el árbol de escenarios considerará 15 etapas, con 1000 unidades de cosecha. La cantidad de escenarios variará en distintas instancias de forma de comprobar las hipótesis planteadas, la generación de estos escenarios se detalla en la siguiente sección.

Para el planteamiento del problema se deben considerar restricciones en cuanto a satisfacer la demanda y por otro lado respecto a limitantes para talar en una misma zona en periodos consecutivos. La demanda es un parámetro que también será estudiado, es decir, se buscará concluir a qué nivel de demanda se puede responder con las condiciones del bosque.

En resumen se tiene:

- Un bosque en la zona central de Portugal, dividido en 1000 unidades de tala, con características similares
- Se generan varias instancias con distinto número de escenarios, y el problema se resolverá para cada una de ellas.
- Cada instancia contará con información de la productividad de las áreas de cosecha.
- Se deben tomar decisiones para 15 periodos anuales, en relación a la proporción que debe ser cosechada en cada periodo para cada unidad. Se puede empezar a cosechar desde el primer periodo.

4.3 Generación de Escenarios

El primer paso a tomar consiste en generar el árbol de escenarios a partir del cual se planteará el modelo de producción forestal. El objetivo del procedimiento aquí descrito es intentar generar un árbol que represente todas o la mayor cantidad de posibilidades de clima futuro.

La información entregada por el programa de la Comunidad Europea corresponde a predicciones mensuales de 143 estaciones meteorológicas, en el periodo comprendido entre 2014 y 2030. Al igual que en el trabajo que lo precede, los escenarios serán basados en ENSEMBLES Project [9] que proveen de bases de datos de clima desarrollados por el Centro Hadley. Estos datos incluyen todos los factores para la productividad de la zona de estudio. En el Anexo B se incluye una muestra de los datos que facilita el programa.

En primer lugar, se agrega la información a nivel anual ya que el modelo se desarrollará con etapas anuales. Luego, un año estará representado por la acumulación de lluvias entre los meses de Marzo y Mayo, que es la época de lluvias.

Por otra parte, se tienen 143 estaciones, con su nivel de precipitación anual correspondiente (según esta definición). Así, se escogen los mínimos y máximos para cada año. En la Figura 4-2 se señala un ejemplo con 5 estaciones meteorológicas y los niveles de precipitación en milímetros de lluvia caída entre Marzo y Mayo.

Figura 4-2 Ejemplo para selección de climas extremos

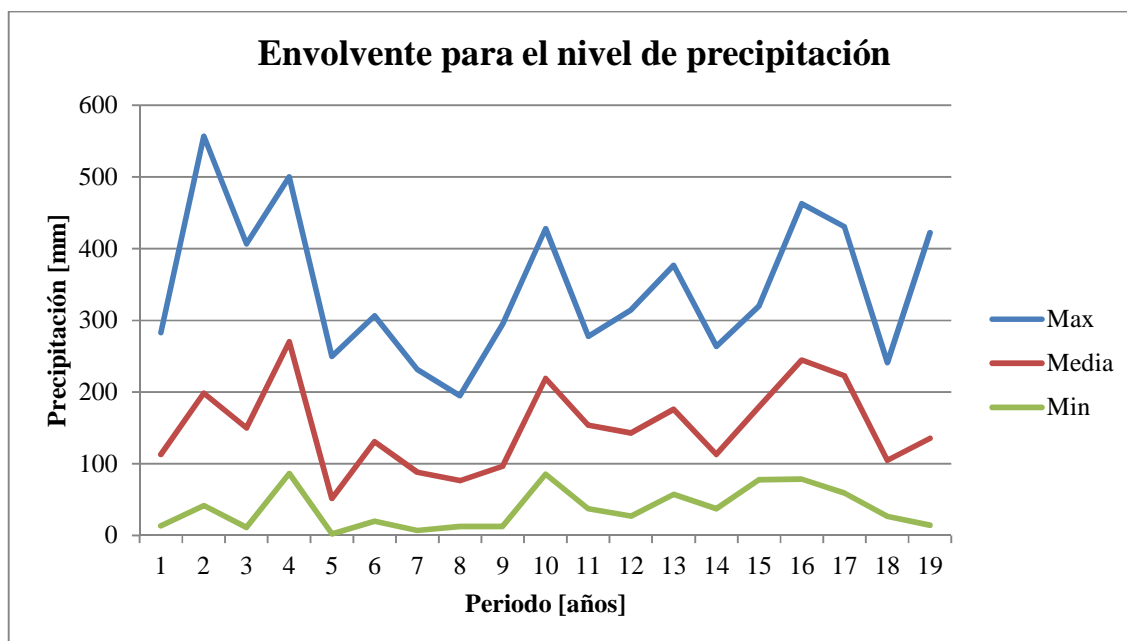
Estación	Año 1	Año 2	Año 3	Año 4	Año 5	Año 6	Año 7	Año 8	Año 9	Año 10	Año 11	Año 12	Año 13	Año 14	Año 15
1	18,8	60	16,3	136,2	4,2	33,7	11,6	13,3	18,7	173,5	55,2	62,6	57,4	49	99,2
2	38,8	66,7	66,8	136,8	11,3	55,5	29,3	28,8	27,2	117,2	104,5	39,3	88,8	62,8	104,4
3	118,2	200,6	154,4	258,3	41,3	136,6	80,9	73,6	79,5	215,5	143,1	143,9	231,4	112,8	170,5
4	115,5	237,8	182,4	301,2	42	142,3	87,2	75,5	94,8	263,1	164,6	185,2	248,2	113,1	220,3
5	260,6	556,4	406,6	500	250	296	231,8	185,1	255	424,6	269,7	313,3	274,9	263,9	296,6

Niveles de precipitación para 5 estaciones meteorológicas, representado como los milímetros de agua caída entre Marzo y Mayo.

En el ejemplo, se tienen 5 estaciones, destacado con color verde están los valores mínimos y con azul los niveles máximos. Los escenarios simulados se moverían en un rango 10% mayor al presentado en la tabla anterior, así se espera estar cubierto frente a cualquier clima que se presente.

En la Figura 4-3 se muestran los valores máximos y mínimos para las 143 estaciones con las que se cuenta información. De dicha envolvente para la precipitación anual, se procede también a extender el rango en un 10%, de forma que efectivamente los escenarios a generar cubran lo más precisamente la totalidad de posibles climas futuros. Estos son los datos que finalmente se utilizarán para simular los escenarios.

Figura 4-3 Niveles mínimos y máximos de precipitación por periodo



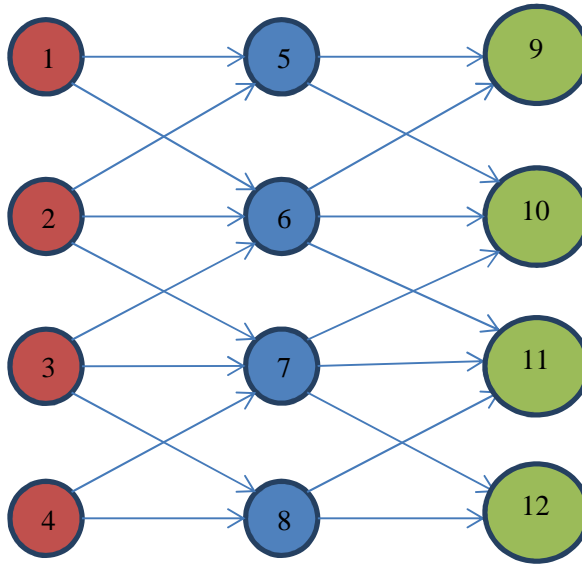
A partir de esta información, se procede a escoger 7 puntos equidistantes, para cada periodo, entre el nivel mínimo y el máximo de precipitación. Estos puntos corresponderán a los nodos del árbol de escenarios final. Para completar la información adicional que se requiere para definir completamente un escenario, se buscará entre las predicciones de todas las estaciones meteorológicas, el año en que la precipitación acumulada en los tres meses elegidos sea la más cercana, y se completarán los otros factores con dichos datos.

Así, se tienen 7 nodos en cada periodo, que deben combinarse de distintas formas periodo a periodo para completar el horizonte de evaluación. Para evitar casos extremos, en que de un periodo a otro el clima sea demasiado disímil, se restringirá a que cada nodo solo se combine con los nodos adyacentes.

En la Figura 4-4 se muestran las posibles combinaciones entre nodos, para tres periodos con solo 4 nodos por periodo. Es así que, por ejemplo, una secuencia factible sería nodos 1 – 6 – 9, pero no es posible pasar del nodo 2 al nodo 8.

Las combinaciones entre los nodos no se mantendrán para todos los periodos, ya que esto generaría una cantidad de escenarios que sería absolutamente imposible de manejar, y en realidad no aporta mayor precisión al modelo. El nodo inicial, de la primera etapa, se combina con los 7 nodos existentes en la segunda etapa. Luego, cada nodo se combinará con los nodos inmediatamente adyacentes, esto sería como en la Figura 4-4, el nodo 2 se combinará con el 5, 6 y 7. Esto se mantiene entre las etapas 1 y 5, es decir, hasta el año 2016, y además entre los periodos 8 y 11, entre los años 2018 y 2021. Para el resto de las etapas, se considerará que continúa en una línea recta (en la Figura 4-4 sería que el nodo 6 combinara solo con el 10).

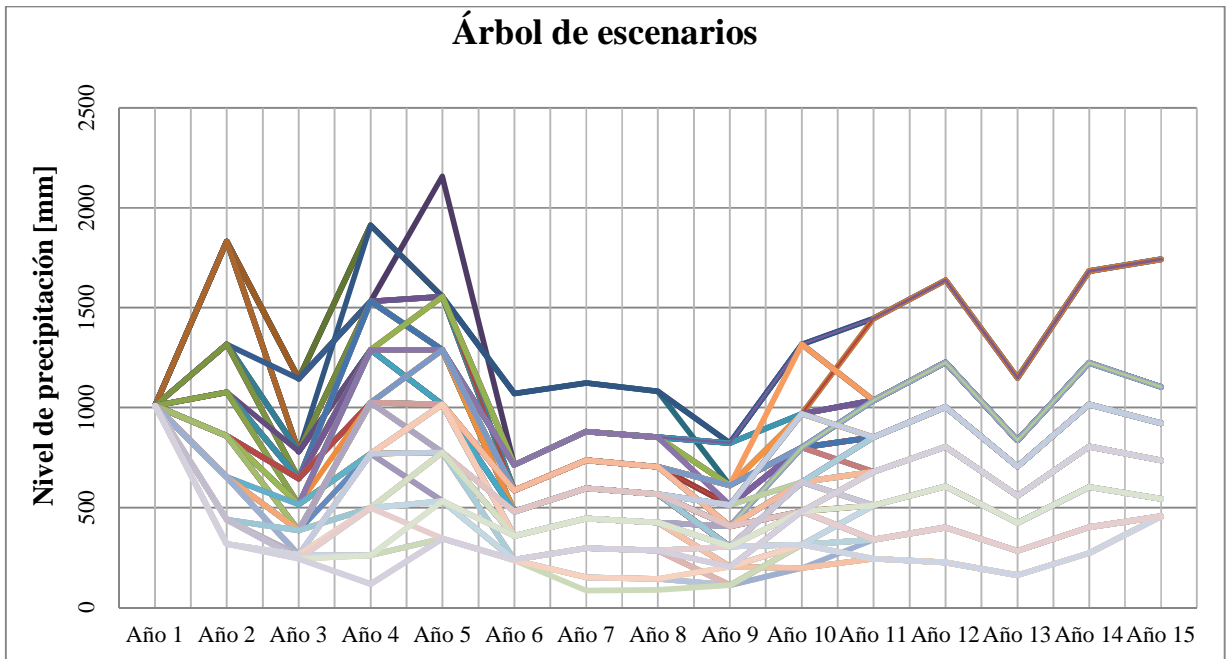
Figura 4-4 Ejemplo de posibles combinaciones entre nodos



Con esta configuración, el número de escenarios generados llegaría a 10062. A partir de este alto valor, se escogen diferentes subconjuntos de escenarios, de forma de contar con diferentes instancias para analizar sus diferentes resultados. La elección de estos subconjuntos se especifica en la siguiente sección.

En el Figura 4-5 se incluye una muestra de 100 escenarios, para representar lo que será a gran escala. Sería imposible graficar los 10062 escenarios generados, pero la estructura es similar.

Figura 4-5 Árbol para una muestra de 100 escenarios



4.4 Reducción del árbol de escenarios

Para llevar a cabo la reducción de escenarios se utiliza una técnica ampliamente estudiada, cuyo punto de partida se centra en realizar una medida (bajo alguna métrica adecuada) de la distribución de probabilidad de la instancia original, con tal de luego llevar a cabo una representación de ésta con un menor número de escenarios, conservando de la mejor manera posible las características inherentes a dicha distribución. Se han desarrollado una serie de algoritmos que abordan este tipo de problema, utilizando diversas métricas para llevar a cabo la medición de las distribuciones de probabilidad y la distancia que separa una instancia de otra, siendo el aporte más destacado el abordado por Heitsch y Römissh [7], donde se utiliza una aplicación de la métrica de Fortet-Mourier para desarrollar dos algoritmos que se encargan de llevar a cabo la selección de aquellos escenarios del set original que mantienen de mejor manera la distribución original.

Los algoritmos más conocidos son fast-forward y back-forward, cuya única diferencia radica en la forma en que se van seleccionando los escenarios. En el primero, se comienza con un conjunto vacío y se van agregando los escenarios seleccionados, mientras que en el segundo se comienza con el conjunto original y se eliminan aquellos que no se desean conservar.

Considerando trabajos anteriores que han utilizado este tipo de técnicas, se trabaja con el algoritmo de fast-forward, que se encarga de calcular la distancia existente entre cada uno de los escenarios que componen la instancia inicial, para luego seleccionar los escenarios que mejor representan la data original.

Una vez elegidos los escenarios que conformarán la selección, se lleva a cabo una redistribución del espacio probabilístico del árbol original, repartiendo las probabilidades de ocurrencia de los escenarios no elegidos al subconjunto seleccionado, utilizando el teorema de Redistribución de probabilidades [7]. De manera simplificada, se agrega la probabilidad de ocurrencia a aquellos escenarios que se encuentren más cercanos. El teorema es presentado en el Anexo A.

La metodología se expone a continuación, de una forma simplificada.

Algoritmo Fast-forward

1. Inicializar: S conjunto original, K número de escenarios a seleccionar.
2. Conjunto objetivo $T = \emptyset$.
3. Mientras $|T| \leq K$:
 - a. Calcular distancia entre escenarios.
 - b. Si $S_i = \operatorname{argmin}\{Métrica\ escogida\}$, entonces $T = T \cup \{S_i\}$.
4. Redistribuir el espacio probabilístico de S en T .

Para utilizar la anterior metodología, es necesario en primer lugar contar con la distribución de probabilidad en cada nodo del árbol. Se concluye que será suficiente, y representativo dada la estructura del árbol, considerar que los nodos hijos sean equiprobables en cada etapa.

El algoritmo Fast-Forward ya ha sido implementado por investigadores que son parte del programa del cual es parte este trabajo de memoria. Se cuenta con el programa que lo implementa y por lo mismo, es directo encontrar los subconjuntos deseados, se indica el número de escenarios

deseados y se obtiene la instancia. Así, de una instancia generada de 10062 escenarios, se reduce dicho árbol para obtener 17 instancias, de 1, 10, 32, 55, 100, 158, 200, 300, 400, 440, 500, 600, 700, 800, 900, 1000 y 1240 escenarios cada una.

Finalmente, cuando se tienen los escenarios seleccionados, será necesario ingresarlos al simulador de productividad entregado por el programa ForEAdapt para obtener el input necesario para encontrar la solución al problema. Es importante destacar que una vez obtenido el resultado del simulador, ya se cuenta con los programas necesarios para transformarlos al formato específico requerido por el software a utilizar para resolver el problema, esto es un programa en JAVA facilitado por otros participantes del programa de la Comunidad Europea y disponible para ser utilizado.

4.5 Modelo de Producción Forestal

Gran parte de la complejidad de este problema radica en las restricciones de no anticipatividad. En la formulación explícita, estas restricciones se definen directamente en el modelo. En una formulación compacta (o implícita), por el otro lado, estas restricciones son expresadas al introducir variables de decisión compartidas en cada nodo del árbol de escenarios. En este trabajo, se utiliza la formulación compacta, mostrada a continuación, ya que lleva a problemas que pueden ser más sencillos de resolver, como indica Quinteros en [10]. En adelante, se referirá a la formulación que sigue como Formulación Extendida, ya que en ningún caso se resolverá la versión explícita.

Conjuntos

$T =$ Horizonte de planificación $\{t\}$.

$H =$ Áreas de tala $\{h\}$.

$S =$ Escenarios $\{s\}$.

Parámetros

$A_h =$ Área de la unidad de tala h

$Pr_s =$ Probabilidad de ocurrencia del escenario s .

$R_{s,t} =$ Precio de venta en periodo t bajo escenario s

$D_{s,t} =$ Demanda mínima a satisfacer en el periodo t bajo escenario s

$a_h^{s,t} =$ Productividad de la unidad h en el periodo t bajo escenario s

$\alpha =$ Fluctuación máxima de volumen permitida entre dos periodos consecutivos

VARIABLES DE DECISIÓN

$x_h^{s,t} =$ Proporción de la unidad h talada en el periodo t bajo el escenario s .

$w_{s,t} =$ Producción total de madera en periodo t bajo escenario s

Se destaca de esta forma de plantear las decisiones, que un área no necesariamente es talada en su totalidad en cierto periodo.

Restricciones

1. *Integridad de la unidad de tala*

$$\sum_{t \in T} A_h x_h^{s,t} = A_h \quad \forall h \in H, s \in S$$

2. *Restricción de tala de una zona en periodos consecutivos*

$$w_{s,t} - (1 + \alpha)w_{s(t+1)} \leq 0 \quad \forall_{s,t}$$

$$w_{s,t} - (1 - \alpha)w_{s(t+1)} \geq 0 \quad \forall_{s,t}$$

3. *Restricción de demanda mínima a satisfacer*

$$D_{s,t} \leq w_{s,t}$$

4. *Naturaleza de las variables*

$$0 \leq x_h^{t,s} \leq 1$$

$$w_{s,t} \in Q^+$$

5. *Producción total*

$$\sum_{h \in H} a_h^{t,s} A_h x_h^{t,s} = w_{s,t} \quad \forall t \in T, s \in S$$

Función Objetivo

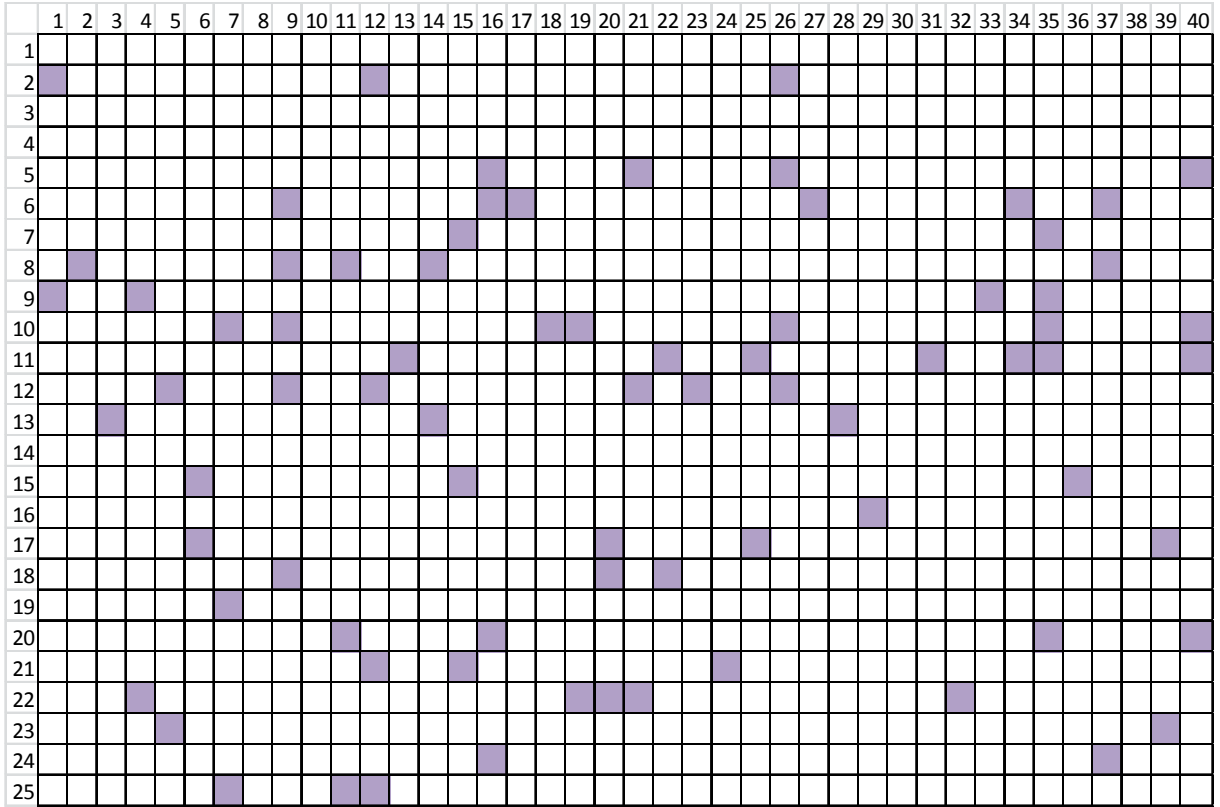
$$\max \sum_{s \in S} \Pr(s) \sum_{t \in T} \sum_{h \in H} R_e^{t,s} a_h^{t,s} A_h x_h^{t,s} \quad \forall s \in S$$

Se maximiza el beneficio considerando el precio de venta para cada periodo, además de la probabilidad de ocurrencia asignada para cada escenario.

El resultado que debería arrojar este modelo, es que para cada nodo del árbol, debiese existir una decisión para cada área de cosecha. Un ejemplo de las decisiones tomadas se muestran en la Figura 4-6, donde se representa el bosque a cosechar como una cuadrícula con 1000 áreas, se grafica una decisión para una instancia de prueba de 55 escenarios, se marcan en color las áreas que deben ser cosechadas en el periodo 10, en uno de los nodos.

La Formulación Extendida hace referencia a que la resolución incluye todos los escenarios de la instancia. En el siguiente capítulo se detalla el algoritmo Progressive Hedging, que a diferencia de resolver el problema para la Formulación Extendida, el algoritmo se caracteriza por resolver a través de la descomposición por escenarios.

Figura 4-6 Representación gráfica de una decisión para instancia de prueba de 55 escenarios, en el periodo 10



Capítulo 5

Algoritmo *Progressive Hedging*

En este capítulo se tratará en detalle el algoritmo *Progressive Hedging*, para lo que en primer lugar se explicitará el procedimiento del mismo, para luego llegar a precisar los distintos ajustes que se pueden hacer para mejorar su eficiencia. Ya existen avances en el estudio de las aplicaciones de este algoritmo en problemas forestales, como los presentados por Pais en [11], sin embargo la configuración debe ser adaptada para las actuales instancias y formulación del problema.

5.1 Pseudo-código extendido

Para facilitar la comprensión del algoritmo y los ajustes que se hacen, se presenta una versión extendida del pseudo-código, incluyendo detalles de la implementación utilizada.

1. $k = 0, g^* = \infty$

2. Para cada escenario $s \in \mathcal{S}$:

$$x_s^k = \operatorname{argmin}_{x_s} f_s(x_s): x_s \in Q_s$$

3. $k = k + 1$

4. Para cada $t \in \mathcal{T}$, para todo nodo $\mathcal{N}_t \in \mathcal{N}$ en el árbol de escenarios:

$$\bar{x}_{n,t}^k = \frac{1}{|\mathcal{N}_t|} \sum_{s \in \mathcal{N}_t} x_{t,s}^k$$

5. $g^k = \sum_{s \in \mathcal{S}, t \in \mathcal{T}} \|x_{t,s}^k - \bar{x}_{n(s,t),t}^k\|$

6. Si se cumplen condiciones, fijar variables.

7. Si $g^k < g^* \vee \sum_{s \in \mathcal{S}} f_s(x_s^k) > \sum_{s \in \mathcal{S}} f_s(x_s^*)$ guardar mejor solución, $x^* = x^k$

8. Si $g^k < \varepsilon \vee k > k_{max}$, ir a 13.

9. Si $k \leq 1, \forall x_s^i \quad \rho_s^i = \rho^i(x, k, s)$

10. $\forall s \in \mathcal{S}, t \in \mathcal{T} \quad w_{s,t}^k = \rho(x_{t,s}^k - \bar{x}_{n(s,t),t}^k) + w_{s,t}^{k-1}, (w^{(0)} = \vec{0})$

11. Para cada escenario $s \in \mathcal{S}$:

$$x_s^k = \operatorname{argmin}_{x_s} f_s(x_s) + \sum_{s \in \mathcal{S}, t \in \mathcal{T}} \left[w_{s,t}^k x_{t,s} + \frac{\rho}{2} \|x_{t,s} - \bar{x}_{n(s,t),t}^k\|^2 \right]: x_s \in Q_s$$

12. Ir al paso 3.

13. Usando x^* como punto de partida con las mismas variables fijadas, resolver la formulación extendida compacta: $\min_x \sum_{s \in \mathcal{S}} f_s(x_s): (x \in Q)$

Consideraciones:

- El paso 2 indica que se resuelve cada escenario por separado.
- En el paso 4 se presentan elementos con la misma probabilidad, por ende, los pesos son idénticos.
- El paso 5 corresponde al cálculo de la convergencia del algoritmo, que es la distancia euclídeana entre cada escenario y su óptimo estimado del nodo correspondiente. También se introduce una función $n: S \times \mathcal{T} \rightarrow \mathcal{N}$ para ubicar a cada escenario con su nodo correspondiente en cada etapa. Así, se tiene una representación simple a la estructura de árbol original con padres e hijos. Del mismo modo, se utiliza n como subíndice para los promedios que presenta un nodo determinado.
- En el paso 7 se adiciona una nueva familia de variables superindexadas con * indicando que son la mejor solución obtenida hasta el momento por el algoritmo. Esto se debe a la naturaleza heurística que no siempre converge o mejora el valor de la función objetivo sino que va saltando de solución en solución.
- En el paso 9 se presenta el penalizador ρ como función dependiente de todas las variables, la iteración y el escenario. Esto varía según el tipo de penalizador utilizado y se trata con detención en la subsección 5.2.3.
- El paso 13 corresponde a una implementación utilizada debido a que llegar al óptimo tomaría demasiado tiempo, se opta por establecer un nivel de gap específico y un límite de tiempo.
- Finalmente, los pasos 1 y 3 son, respectivamente, la inicialización y la actualización del iterador. El paso 6, el fijado de variables, se trata con detención en la subsección 5.2.5.

5.2 Ajustes y mejoras

Las opciones investigadas son:

- Acelerar la resolución de cada subproblema.
- Linealizar el término cuadrático proximal.
- Ajustar el parámetro ρ de penalización.
- Paralelizar el algoritmo.
- Fijar variables.
- Variar el criterio de término.
- Clustering de escenarios
- Bundling de escenarios

5.2.1 Acelerar la resolución de cada subproblema

El ajuste del gap a optimalidad (también conocido como mipgap) es fundamental para reducir el tiempo de resolución del algoritmo, ya que encontrar el óptimo en este problema puede tomar horas, mientras que una solución buena puede tardar algunos minutos. Por lo tanto, es muy

conveniente que no sea necesario encontrar el óptimo de cada subproblema para que el algoritmo converja. Así, la técnica consiste en encontrar un mipgap lo suficientemente grande como para que cada subproblema tome poco tiempo, pero a la vez lo necesariamente pequeño para que las soluciones sean factibles.

Finalmente, luego de varios ensayos, se optó por un gap de 2%, que cumple las condiciones necesarias en cuanto al tiempo de resolución y la convergencia del algoritmo.

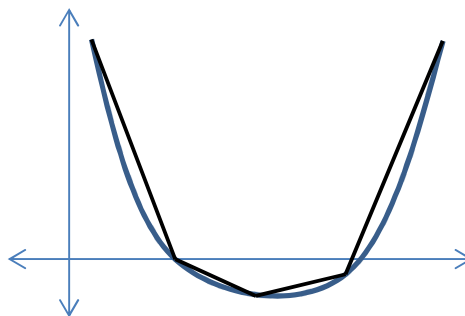
5.2.2 Linealizar el término cuadrático proximal

La única gran desventaja de utilizar el algoritmo Progressive Hedging es que al resolver cada subproblema, a pesar de que son simples, es necesario agregar el término proximal cuadrático de la penalización (paso 6 del algoritmo). Esto complica bastante cada problema y puede aumentar el tiempo de solución.

$$\frac{\rho}{2} \|X^k(s) - \bar{X}^{(k-1)}\|^2$$

Este término es necesario para la convergencia de la heurística, logrando que cada variable de cada escenario se acerque al promedio del nodo respectivo. La técnica a utilizar consiste en reemplazar este término por algo parecido, en este caso, se puede hacer una aproximación lineal por tramos como lo que se muestra en la Figura 5-1.

Figura 5-1 Diagrama de linealización por tramos



En la práctica, el comportamiento de esta aproximación depende de la cantidad de segmentos escogidos. Cada segmento se considera como una variable nueva. El número de segmentos es proporcional al número de vértices del poliedro que representa el subproblema. Luego, si se tienen pocos segmentos, como en el caso extremo con 2 segmentos, entonces la penalización queda con forma de V, y así todas las variables “caen” al promedio, el problema converge, pero con una mala solución en relación a la función objetivo. Pero por otra parte, al tener muchos segmentos, aumenta el número de variables y se vuelve a complicar la resolución, que es precisamente lo que se quiere evitar.

En base a lo anteriormente planteado, es lógico notar que existe un trade-off entre el número de tramos a elegir para realizar la linealización, y por ende, la velocidad de resolución de cada subproblema, y la calidad de la solución obtenida. Sin embargo, esta situación puede contrarrestarse a partir de elegir un tamaño adecuadamente pequeño para el penalizador ρ , elemento que equipara el efecto negativo de elegir una cantidad disminuida de segmentos lineales dentro de la aproximación, ya que esto permite que muy pocas variables se vean forzadas a

converger en cada iteración, pero muy velozmente, sin empeorar así el valor de la función objetivo obtenida.

Por lo tanto, se considera tomar la máxima linealización, es decir dos tramos, para luego contrarrestar el efecto escogiendo sabiamente el parámetro ρ como se señala en la siguiente sección.

5.2.3 Ajustar el parámetro ρ de penalización

El parámetro ρ es el ponderador relativo entre forzar la no anticipatividad y cumplir el término proximal. También por la construcción del ponderador w , ρ se puede interpretar como el precio fijado por desviarse de la solución ponderada y así w tiene la interpretación de ser el cobro extra dentro de la función objetivo de manera que para el subproblema sea mejor tomar la solución que cumpla la no anticipatividad.

Luego, el beneficio viene dado al ajustar este parámetro para encontrar un sistema de precios que mejore la relación de las variables. De esta forma el algoritmo llegará a converger más rápido y con un rendimiento superior en términos de función objetivo y calidad.

Las estrategias testeadas y estudiadas durante el transcurso de este trabajo consisten en tratar con un parámetro ρ fijo y con la posibilidad de que éste varíe en las distintas iteraciones. Junto a lo anterior, el método para seleccionar el valor de ρ se desarrolla bajo el anhelo de evitar una convergencia precoz de todas las variables producto de la linealización extrema en forma de embudo (solo 2 tramos), la rapidez de cómputo en cada iteración y la velocidad de resolución del problema completo.

El estudio de un parámetro variable no tuvo resultados positivos, en particular porque para gran parte de las instancias, se pudo encontrar buenas soluciones en las primeras iteraciones por lo tanto no vale la pena variar ρ . Tomando un parámetro $\rho = 0,0001$ se obtuvieron los mejores tiempo de resolución para la mayor parte de las instancias.

5.2.4 Paralelizar el algoritmo

Debido a sus características intrínsecas a la hora de tratar el problema original y generar una descomposición de éste según escenarios (es decir, descomposición horizontal de la instancia original) es que el algoritmo de Progressive Hedging, presenta una estructura natural para llevar a cabo programación paralela y distributiva, enviando cada uno de los subproblemas (pequeños y de veloz resolución) a un núcleo o nodo de un clúster determinado, obtener la solución de éste y luego hacer concurrir las soluciones para actualizar el estado de los parámetros del algoritmo y continuar a la siguiente iteración.

Su implementación computacional para el algoritmo PH se encuentra en el software PYSP, parte del paquete de software Coopr de los Laboratorios Nacionales Sandia, pudiéndose así realizar una serie de experimentos en relación al algoritmo y el número de procesadores que participan de su resolución. Los resultados obtenidos van de mano con lo esperado, puesto que los tiempos de resolución de las instancias experimentales mejoraron sustancialmente, hasta un 70%, en comparación al caso serial y al de formulación extendida. Las mejoras en el tiempo se presentan fuertemente en el caso de las instancias con mayor número de escenarios. Así, se

utilizarán en cada caso 4 núcleos, una cantidad que tienen las máquinas accesibles a todo público en la actualidad.

El esquema de trabajo utilizado por esta metodología consiste en la presencia de un computador maestro dentro del cual se encuentran los modelos e instancias a tratar, el cual actúa simplemente como monitor de una serie de procesadores, que se encargan de resolver cada uno de los subproblemas que éste envía. El computador maestro establece un servidor virtual que se encarga de gestionar la conexión con el resto de los “solucionadores” a través de un servidor de despacho, que se encarga de enviar y recibir los subproblemas y sus respectivas soluciones para poder avanzar a través de las iteraciones. Cada uno de los nodos o procesadores establece conexión a este servidor virtual a través de un servidor de resolución (mip server), que presenta una dirección única e inequívoca con dicho recurso computacional. Todo lo anteriormente mencionado fue implementado en una serie de scripts programados en Python y MS_DOS, complementarios al paquete PYSP, de tal forma de poder utilizar esta técnica de forma automática y directa, simplemente determinando de forma previa el número de procesadores que se pretenden utilizar.

Importante es notar que el script utilizado busca distribuir la carga de trabajo sobre cada uno de los procesadores o nodos “solucionadores” de la forma más equitativa posible, de tal modo que en promedio se tengan todas las soluciones de los escenarios dentro de una ventana de tiempo establecida.

5.2.5 Fijar variables

Esta estrategia toma como punto de partida lo planteado por Watson y Woodruff en [8] donde la idea central es que para acelerar la convergencia de la heurística, es necesario fijar las decisiones que ya estén convergidas (iguales en los escenarios que comparten un nodo respectivo dadas las restricciones de no anticipatividad) para tratar con subproblemas más pequeños y aumentar la velocidad de resolución.

La necesidad de esta técnica radica en que a lo largo del desarrollo de la heurística las variables enteras no consiguen alcanzar una estabilidad en cuanto a valor, por lo que comienzan a experimentar un “vaivén” o “saltos” entre las diversas soluciones, lo que impide que alcancen la convergencia de forma óptima. Por ende, la estrategia consiste en ir fijando aquellas decisiones que según los escenarios que comparten nodos de información respectivos, hayan alcanzado un estado estable, asumiendo así que dichas decisiones son las definitivas, supuesto que si bien es fuerte, puede mejorarse aplicando una serie de técnicas complementarias. Una de éstas consiste en esperar a que las variables alcancen un estado estacionario de convergencia y que no cambien durante un número determinado de iteraciones, para luego, en caso de alcanzarse dicho número, fijarlas.

Las estrategias utilizadas son:

- **Fijar tras la iteración cero:** corresponde a la estrategia donde se fijan aquellas decisiones que alcanzan la convergencia la primera vez que se resuelve cada uno de los escenarios por separado, por ende se observa que si aun sin considerar la no anticipatividad existen ciertas decisiones que la cumplen entonces deben mantenerse.

Claramente el supuesto detrás de esto es bastante fuerte, ya que asume que la solución final que se alcanzará se encuentra compuesta por la solución inicial obtenida, hecho que puede ser totalmente erróneo en algunos casos.

- **Fijar con retardo:** Consiste en fijar las decisiones convergidas y que no hayan cambiado luego de un cierto número de iteraciones o retardo. Presenta una mayor solidez en comparación a la estrategia anterior, debido a que no se realiza un supuesto tan fuerte. Su implementación se realiza de tal manera que es posible distinguir el número de retardos según la etapa que se está resolviendo.

De los experimentos, se concluye que la mejor configuración consiste en fijar variables después de la primera iteración. Fijar en la iteración cero tiene un supuesto demasiado fuerte y empeora la calidad de la solución, mientras que probando con distintas instancias, fijar en la primera iteración permite mejorar los tiempos sin perder calidad.

5.2.6 Variar el criterio de término

El criterio de término del algoritmo, si bien incluye terminar luego de alcanzar la tolerancia mínima para aceptar la convergencia, en la práctica esto no se da puesto que toma mucho tiempo y el valor de la función objetivo empeora al forzar que el problema converja. Luego el criterio que se usó fue iterar un número razonable veces de tal manera que se fijen suficientes variables para lograr resolver la formulación extendida compacta con variables fijadas de manera rápida.

El criterio obtenido empíricamente fue un valor para $\varepsilon = 0,01$, mientras que se mantuvo un número de 5 iteraciones como máximo.

5.2.7 Clustering de escenarios

Al momento de abordar problemas de programación estocástica se debe tener en consideración la gran dimensionalidad de los árboles con los que se puede llegar a trabajar y por lo mismo es importante diseñar estrategias que permitan representar de la manera más fiel posible las posibilidades pero que a la vez sea un problema que sea factible de ser resuelto y adicionalmente en un tiempo y con recursos computacionales adecuados. En este espíritu, es que se puede considerar no utilizar todos los escenarios iniciales, sino que una muestra representativa de éstos de forma de mantener la distribución de los datos originales y que por lo tanto los resultados se vean afectados en la menor medida posible.

Se deben considerar algunos requisitos esenciales para no alterar la validez teórica de la solución y que también sea comparable al resultado que se obtendría al utilizar la totalidad de la instancia original. Estos requisitos principalmente son que se mantengan los casos extremos, es decir, los escenarios en donde el clima presenta la mayor y la menor cantidad de precipitación, y por otra parte, que el subconjunto de escenarios escogidos represente de forma satisfactoria la estocasticidad del problema original.

La alternativa que surge de forma natural consiste en la confección de Clusters, pues es una aplicación sencilla y que ofrece una primera aproximación natural al problema de generar un subconjunto representativo de un set de datos original. La idea consiste en generar una serie de clusters o grupos que en la práctica representan un conjunto de escenarios que comparten características similares en cuanto al comportamiento de sus variables aleatorias, logrando así

resumir la información contenida en un gran número de escenarios en un solo vector. Sin embargo, uno de los problemas que presenta esta técnica de reducción de escenarios consiste en que presenta un sesgo a mantener escenarios promedios, dejando de lado eventos extremos donde las variables que representan la estocasticidad del problema puedan dispararse o caer abruptamente, lo que pese a tener una baja probabilidad de ocurrencia, puede llevar a grandes pérdidas si es que no son tomados en cuenta a la hora de procesar el modelo.

Con tal de hacer frente a la situación mencionada, es posible forzar a los algoritmos para construir los clusters a mantener un determinado número de eventos extremos de la instancia original, conservando una baja probabilidad de ocurrencia.

Para la confección de los clusters, se utilizaron 2 técnicas. En primer lugar, se tomó el algoritmo de conglomerados de K medias, con una modificación de forma de incluir los climas más extremos, y por otro lado, se construyen clusters de tamaño uniforme. Ambas técnicas son explicadas a continuación.

Del mismo modo, es preciso señalar que para mantener la validez de los escenarios generados, no se debe utilizar el valor de un “escenario promedio” como típicamente se realiza al utilizar esta técnica, sino que se procederá a calcular los diferentes clusters que representan de mejor manera el conjunto de escenarios original, y luego, se determinará el escenario más cercano al promedio de cada clúster, conservándose así los escenarios que cumplan con esta última condición. Así, se evita trabajar con medias de los datos originales, manteniendo la distribución y comportamiento de la instancia creada.

Gracias a esta técnica, es posible hacer frente al crecimiento explosivo de escenarios que se produce en la medida que se trabaja con un mayor horizonte de planificación y/o se aumenta el número de nodos hijos por cada nodo según etapa.

Conglomerado de K medias

Se utiliza una adaptación específica del algoritmo de clustering k-means clásico, algoritmo que recibe como input el número de clusters que se pretenden encontrar dentro del set de datos original, a través de la búsqueda de los K puntos medios representativos de dicho conjunto.

Este algoritmo fue modificado especialmente para este trabajo de tal forma de que se mantengan en todo momento los eventos extremos del conjunto original, se utiliza como criterio mantener las “colas” del árbol de escenarios, así se obliga a mantener el 5% de los escenarios con el clima más seco y el 5% de los que tienen el clima más lluvioso.

El procedimiento utilizado es el siguiente:

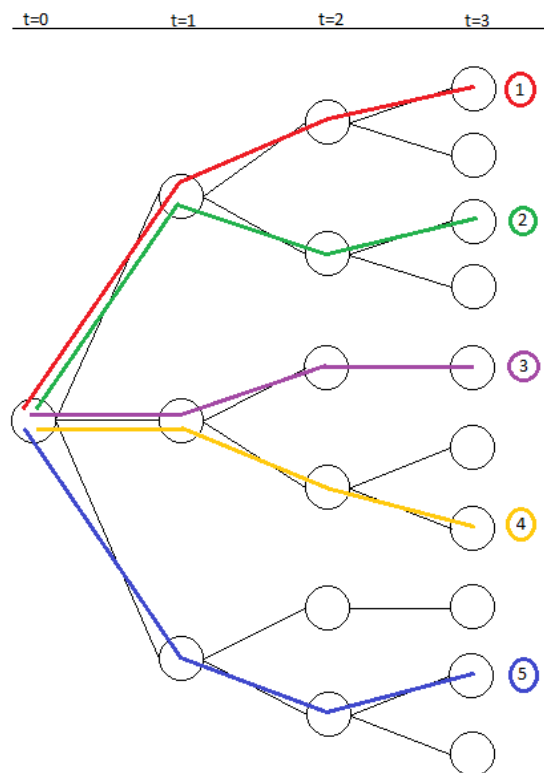
1. Se inicia con el conjunto original de escenarios S.
2. Se crean K clusters, cada uno de ellos identificado con un centroide.
3. Se identifican los escenarios pertenecientes a cada cluster.
4. Se escoge por cada cluster, el escenario del conjunto original S más cercano al centroide, y se incluye en el subconjunto final escogido.
5. Se identifican las “colas” y también se incluyen en el subconjunto final.

Clusters de tamaño uniforme

Se seleccionan escenarios como centros de los clusters de manera uniforme como se muestra en la Figura 5-2. Esto se basa en la idea de que los escenarios al encontrarse ya ordenados dentro del árbol desde los climas más áridos a los más lluviosos, seleccionar de esta forma llevará a cubrir todo el espectro de posibilidades.

En el ejemplo, se muestra cómo se seleccionaría el 50% de los escenarios usando esta técnica, resultando un árbol que es bastante representativo tomando en cuenta la distribución original.

Figura 5-2 Ejemplo de clusters de tamaño uniforme



Se realizarán dos experimentos para cada técnica, variando el número de clusters a crear, de esta forma se analizará el efecto que tiene sobre el resultado final. Se escogen 25% y 50% de los datos originales, siguiendo los procedimientos anteriores.

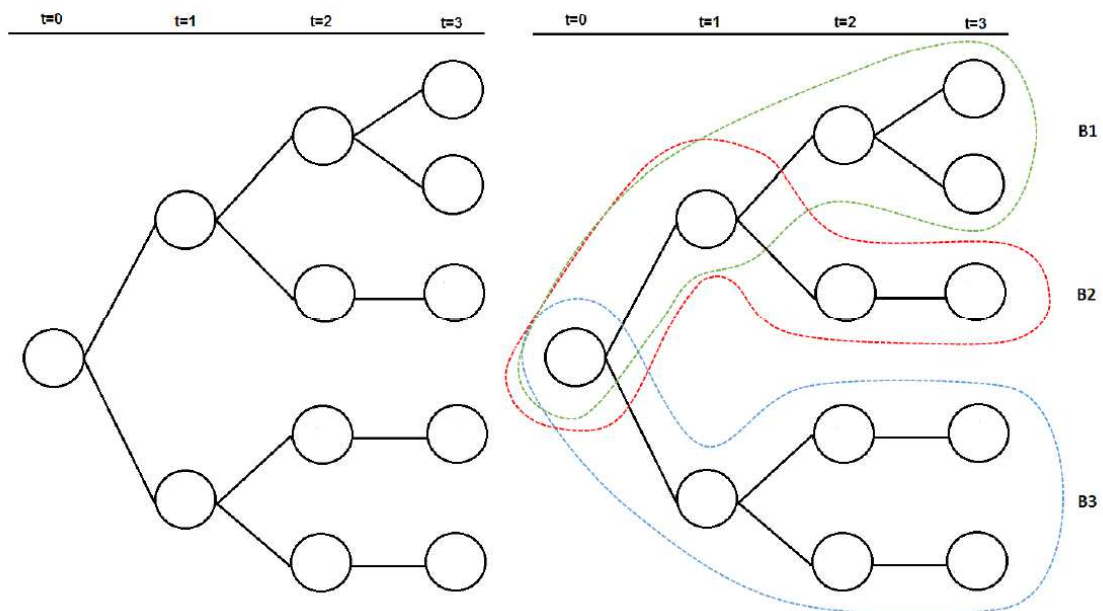
5.2.8 Bundling de escenarios

El concepto de bundling de escenarios hace referencia a una técnica donde se trabaja con conjuntos de escenarios que al menos comparten un conjunto de información del árbol de escenarios original. Se presentan así una serie de sub-árboles de escenarios que se intersectan en determinados nodos a lo largo de su realización. A diferencia de la técnica de clustering, donde se reduce la cantidad de escenarios, en el caso de bundling se toman en cuenta todos los posibles escenarios, pero agrupados en “bundles”.

Así, se persigue trabajar con estos bundles o “macro-escenarios” en lugar de una serie de escenarios simples, lo que permite reducir el tamaño de la instancia de trabajo, resolviéndose un número menor de sub-problemas, sin embargo, estos serán claramente más complejos que los originales, por ende, existe un trade-off entre el número de escenarios que se busca reunir como máximo dentro de un único bundle y el rendimiento a la hora de resolver la instancia, puesto que el resultado podría ser contraproducente, dado que se trabaja con un número menor de sub-problemas, pero más complejos que los originales, ya que cuentan con mayor cantidad de variables y restricciones.

En la Figura 5-3 se representa de manera gráfica la construcción de tres bundles para un árbol que originalmente tiene 5 escenarios.

Figura 5-3 Ejemplo de bundling de escenarios



El objetivo final de este enfoque consiste en que si los macro-escenarios se resuelven de forma más veloz que el tiempo que se toma el algoritmo PH clásico, será conveniente reducir el número de veces que se resuelve cada escenario. Así, en el caso original se resuelven muchas iteraciones de problemas sencillos, mientras que en el caso bajo el enfoque de bundling, se buscará resolver menos iteraciones, en las que los conjuntos de escenarios son más complejos y requieren un mayor tiempo de resolución.

Para su confección, se utilizan dos enfoques en el presente trabajo de tesis, correspondientes a una selección aleatoria y otra basada en la técnica de clustering k-means, donde se realizan bundlings con aquellos escenarios que forman parte de un mismo cluster, pues así se apunta a que presentan características similares y por ende su resolución no debiese complejizarse en gran medida.

Capítulo 6

Implementación

La implementación computacional se realizó utilizando principalmente el software Coopr/PYSP de los laboratorios Sandia [4], consistente en un paquete de software específico para la resolución de problemas de optimización estocástica.

El entorno de trabajo se ubica bajo el lenguaje de programación Python, siendo PYOMO el paquete principal que permite llevar a cabo el modelamiento y resolución de los modelos estocásticos en estudio, permitiendo una definición natural del problema determinístico y su extensión a una estructura estocástica representada por un árbol de escenarios donde se contenga la información relativa a las posibles realizaciones futuras de las variables aleatorias incluidas. Así, es directo llevar a cabo varias series de experimentos, utilizando la implementación de la generación de escenarios, posteriormente la selección de escenarios para cada instancia y finalmente la implementación en JAVA de un programa que transforma el output del simulador de productividad entregado por ForEAdapt en un archivo con el formato requerido por el paquete de modelamiento para resolver el árbol.

El solver utilizado para correr todas las instancias corresponde a CPLEX 12.6, el cual es llamado a través de una serie de scripts programados en Python 2.7, así como de archivos de batería MS-DOS en Windows versión 8.1.

Todas las instancias utilizan la misma extensión del bosque y sus parámetros en su totalidad, siendo la única diferencia el total de escenarios a procesar por cada una de éstas.

6.1 Formulación extendida: PYOMO

La implementación serial del problema en estudio se realizó bajo el paquete de modelamiento PYOMO, desarrollándose todos los modelos testeados en su totalidad bajo este formato, a través del lenguaje de programación Python.

Junto a lo anterior, se utiliza el script “runef” del paquete PYSP de los laboratorios Sandia, script desarrollado en Python que permite crear la formulación extendida de una instancia a partir de la estructura de un árbol de escenarios clásico en formato PYSP, así, su implementación fue directa desde el programa en JAVA.

Del mismo modo, todas las instancias trabajadas fueron luego generadas en formato LP, cuya sintaxis es la utilizada por el solver CPLEX para la lectura y generación de modelos, lo que permite una mayor velocidad de lectura y resolución de los problemas en estudio.

6.2 Progressive Hedging: PYOMO - PYSP

La implementación en paralelo del algoritmo de Progressive Hedging se realizó utilizando el script “runph” del paquete de software Coopr de los laboratorios Nacionales de Sandia, script que permite una implementación directa y natural con el paquete “pickle” de Python [12], el que

permite establecer la creación de una serie de servidores locales dentro del computador o estación de trabajo multi-núcleo utilizada.

Así es posible utilizar programación distributiva dentro de un computador con varios núcleos, enviando cada uno de los subproblemas a distintos puntos de cómputo y estableciendo colas ad-hoc según la carga de trabajo que experimente cada uno de éstos, y luego generar la comunicación entre ellos para continuar con la siguiente iteración del algoritmo.

Un trabajo que quedaría pendiente de corroborar, es utilizar una serie de estaciones de trabajo (típicamente distribuidas a modo de nodos donde se encuentran cúmulos de procesadores). Esto podría permitir una mejora en el rendimiento del algoritmo, ya que en muchos casos al utilizar un mayor número de núcleos en una misma máquina, ésta depende de la memoria RAM disponible, y al verse agotada comienza a recalentarse, por lo tanto pierde eficiencia y empeora considerablemente el tiempo de resolución.

Del mismo modo, permite una implementación directa en su código fuente de las técnicas de mejora explicadas en la sección 5.2, tales como el fijado de variables binarias, variación de mipgap y cálculo del parámetro ρ según iteración. Junto a lo anterior, presenta la posibilidad de añadir propios módulos de programación por parte de usuarios avanzados o plug-ins, lo que conlleva a un manejo profundo de todos los parámetros (número de bundles y la forma de determinarlos, por ejemplo), customizando de forma completa la implementación del algoritmo, lo que otorga así un sin límite de posibilidades para el usuario (una solución a la medida).

6.3 Calidad de la solución

La implementación de esta metodología se realiza en una rutina de JAVA, para cada una de las instancias en comparación a la instancia mayor, de 1240 escenarios.

Primero, se toma la solución encontrada para cada conjunto de datos (10, 32, 55, 100, 158, 200, 300, 400, 440, 500, 600, 700, 800, 90 y 1000 escenarios), luego la rutina identifica las decisiones tomadas para el escenario más cercano a cada uno de los 1240 originales. Si se tiene, por ejemplo, la instancia de 100 escenarios, se encontrará para los 1240, cuál de esos 100 es más similar en cuanto a las condiciones climáticas. Se podría entender que se agrupan los 1240 escenarios en 100 conjuntos, donde los centros de esos conjuntos son los 100 escenarios de la instancia pequeña.

Luego, se comprueba la factibilidad de la solución encontrada para las instancias pequeñas, en las condiciones dadas por el escenario específico de la instancia mayor. Siguiendo con el ejemplo, se entiende que en ciertos escenarios de los 1240 podría darse que ninguna de las soluciones encontradas para los 100 escenarios satisfaga las restricciones del modelo.

Un escenario puede resultar infactible por múltiples razones: no cumplir con las restricciones de flujo, no poder satisfacer la demanda, entre otras cosas. Es decir, sería infactible si al tomar las decisiones que indica la solución al escenario más cercano (qué parte de cada área de cosecha talar en cada periodo), no se cumplen todas las restricciones del modelo.

Para los casos de reducción de escenarios por medio de la técnica de clustering, también se realiza un test de factibilidad, pero esta vez en relación a la instancia a partir de la cual se reduce. Por ejemplo, teniendo la instancia de 100 escenarios, se evalúa qué parte de ese total es factible cuando se encuentra la solución para un subconjunto, por ejemplo, de 50 escenarios.

6.4 Soporte computacional

Las máquinas utilizadas para llevar a cabo los experimentos del presente trabajo de memoria corresponden a computadores de simple acceso para el público. Se trabajó con dos equipos, de forma de realizar distintas pruebas de forma simultánea, teniendo distintas capacidades: el computador 1 trabajó con las instancias más pequeñas y el computador 2 con las instancias masivas. Las características de cada computador son las que se muestran a continuación:

- **Computador 1**

Sistema Operativo: Windows 8.1

Memoria RAM: 12 GB

CPU: Intel Core i5-3317U 1.70 GHz

Solver: CPLEX 12.6 64 bits

Scripts: Python 2.7 64 bits

Middleware: Sandia Labs Coopr Package 3.4.7842

- **Computador 2**

Sistema Operativo: Windows 8.1

Memoria RAM: 16 GB

CPU: Intel Core i7 4700 2.4 GHZ

Solver: CPLEX 12.6 64 bits

Scripts: Python 2.7 64 bits

Middleware: Sandia Labs Coopr Package 3.4.7842

Capítulo 7

Resultados

En este capítulo se presentan los resultados de este trabajo y la metodología antes presentada. En primer lugar, se estudian los resultados, tanto de la implementación extendida como de Progressive Hedging, para distintos niveles de demanda. Posteriormente, se busca realizar una comparación entre los valores de la función objetivo, gap y tiempo de resolución entregado por la formulación extendida, la cual es resuelta directamente en CPLEX, con respecto a la resolución de la misma instancia bajo una implementación del algoritmo PH. Luego, se muestran los resultados para las técnicas de reducción y agrupación de escenarios. Finalmente, se hace un análisis de factibilidad y se concluye sobre la calidad de las soluciones.

Las instancias de prueba constan de 1, 10, 32, 55, 100, 158, 200, 300, 400, 440, 500, 600, 700, 800, 900, 1000 y 1240 escenarios. Todas las instancias incluyen el total del bosque en estudio, variando únicamente el número de escenarios a considerar. Se considera el caso en que exista solo un escenario, equivalente al promedio de clima, y que sería la solución de un modelo determinístico.

La metodología para llevar a cabo los experimentos incluye en la reproducción automática de cada una de las instancias mediante el uso de scripts, los cuales se encargan de llamar a los paquetes involucrados en cada experimento. Luego, los resultados son almacenados en un formato específico, en carpetas identificadas para cada instancia y con la información necesaria sobre el número de escenarios y las diferentes características de cada “corrida”. El objetivo de esto es automatizar y evitar posibles errores en los resultados.

7.1 Resultados generales

Para todas las instancias, se probaron distintos niveles de demanda, de forma de corroborar qué valor de la misma es factible bajo las condiciones actuales, y la que se mantendrá en los experimentos siguientes. Los resultados para los distintos niveles de demanda para instancias de 1, 10, 200, 500, 800 y 1240 escenarios se muestran en la Figura 7-1. Los resultados para todas las instancias de prueba se encuentran en el Anexo C. Se muestran los resultados tanto para la formulación extendida como para PH, de forma de ejemplificar que el comportamiento al variar la demanda es común para ambas formulaciones.

Se puede observar la tendencia según lo esperado, donde a mayor demanda, aumenta la complejidad del problema y por lo tanto los tiempos de resolución. Esto se hace cada vez más marcado al observar las instancias de mayor tamaño. Se decide trabajar con una demanda de 80.000 m³, de forma de encontrar soluciones para el problema más completo. Por otra parte, trabajar con niveles bajos de demanda no guarda demasiada relación con la realidad, ya que la mantención de un bosque de las características del área de estudio es viable solo en el caso de que las demandas por la cosecha sean altas.

Figura 7-1 Resultados para distintos niveles de demanda para instancias de 1, 10, 200, 500, 800 y 1240 escenarios

		Formulación extendida		Progressive Hedging				Formulación extendida		Progressive Hedging	
N° Esc	Demanda	Tiempo [s]	Gap [%]	Tiempo [s]	Gap [%]	N° Esc	Demanda	Tiempo [s]	Gap [%]	Tiempo [s]	Gap [%]
1	0	60,1	0,01%	60,1	0,01%	500	0	11874	1,04%	9974	1,24%
	10000	61,2	0,01%	61,2	0,01%		10000	12746	1,09%	10004	1,22%
	20000	54,7	0,01%	54,7	0,01%		20000	15789	1,19%	10115	1,04%
	30000	59,7	0,01%	59,7	0,01%		30000	18774	1,17%	10875	0,98%
	40000	55	0,01%	55	0,01%		40000	20498	0,98%	11365	0,97%
	50000	53,2	0,01%	53,2	0,01%		50000	24169	0,74%	11497	1,57%
	60000	59,4	0,01%	59,4	0,01%		60000	26789	1,54%	11999	1,69%
	70000	67,2	0,01%	67,2	0,01%		70000	27998	2%	12785	1,87%
	80000	74,2	0,01%	74,2	0,01%		80000	28996,1	1,40%	13478	1,44%
10	0	154	0,01%	204	0,01%	800	0	NO SOL	NO SOL	20141	1,77%
	10000	159	0,01%	209	0,01%		10000	NO SOL	NO SOL	21741	1,11%
	20000	201	0,01%	215	0,01%		20000	NO SOL	NO SOL	21984	1,24%
	30000	213	0,01%	236	0,01%		30000	NO SOL	NO SOL	21657	1,74%
	40000	224	0,01%	227	0,01%		40000	NO SOL	NO SOL	22436	1,65%
	50000	251	0,01%	289	0,01%		50000	NO SOL	NO SOL	23006	1,78%
	60000	274	0,01%	324	0,01%		60000	NO SOL	NO SOL	24789	1,45%
	70000	268	0,01%	333	0,01%		70000	NO SOL	NO SOL	27548	1,98%
	80000	301	0,01%	389	0,01%		80000	NO SOL	NO SOL	29545,2	1,89%
200	0	1040	0,32%	2789	0,40%	1240	0	NO SOL	NO SOL	30741	2%
	10000	1099	0,34%	2478	0,33%		10000	NO SOL	NO SOL	35987	2%
	20000	1178	0,14%	2984	0,41%		20000	NO SOL	NO SOL	36581	1,89%
	30000	1278	0,59%	3174	0,57%		30000	NO SOL	NO SOL	36665	1,88%
	40000	1394	0,47%	3299	0,59%		40000	NO SOL	NO SOL	37884	1,78%
	50000	1662	0,24%	3487	0,40%		50000	NO SOL	NO SOL	38456	1,74%
	60000	1802	0,78%	3399	0,30%		60000	NO SOL	NO SOL	39991	2%
	70000	1877	0,98%	3719	0,74%		70000	NO SOL	NO SOL	44251	1,98%
	80000	1947	0,23%	3879	0,650%		80000	NO SOL	NO SOL	43200	3,12%

Por otra parte, de los experimentos iniciales, y como se expresó en la sección 5.2, la mejor configuración encontrada para el algoritmo Progressive Hedging es con un punto de corte de $\varepsilon = 0,01$, con un máximo de 5 iteraciones. Se fijan variables luego de la primera iteración, el valor para el factor de penalización es $\rho = 0,0001$, linealización del término proximal en 2 tramos y se paraleliza el algoritmo en 4 núcleos. Esta configuración es la utilizada para todos los experimentos posteriores.

7.2 Comparación entre Modelo Determinístico y Estocástico

Al utilizar técnicas de optimización estocástica, que por consiguiente involucran el uso de muchos recursos computacionales, es crucial justificar el uso de muchos escenarios, y esto es principalmente debido a que la solución del modelo estocástico es más robusta frente a la

incertidumbre que el determinístico. Esto fue demostrado en el trabajo que antecede a esta memoria, y se busca reafirmar este concepto.

Una forma de cuantificar los beneficios obtenidos al utilizar distintos escenarios de clima, es analizar las decisiones tomadas en el escenario promedio (correspondiente al modelo determinístico) en las condiciones de los escenarios extremos.

En la Figura 7-2 se muestran los valores de la función objetivo, es decir el beneficio obtenido, para los 2 escenarios más extremos de la instancia de 158 escenarios, estos son el clima más lluvioso y el más seco. Se muestran, por un lado, los resultados implementando las decisiones del modelo determinístico y por otro lado, las decisiones del plan contingente del modelo estocástico (las decisiones específicas para cada uno de esos escenarios).

Figura 7-2 Comparación entre Modelo Determinístico y Estocástico

Escenario	Beneficio Obtenido (€)		Ganancia
	Determinístico	Estocástico	
Lluvioso	\$ 60.333.413,14	\$ 69.117.543,14	\$ 8.784.130,00
Seco	\$ 36.204.792,36	\$ 41.789.583,12	\$ 5.584.790,76

Adicionalmente, es importante destacar que la solución del modelo determinístico no es factible para ninguno de los escenarios escogidos para esta muestra. Es decir, tomando dichas decisiones, no es posible satisfacer todas las restricciones del modelo, en este caso no se satisfacen las restricciones de demanda ni de flujo.

Se observa de los resultados que el modelo estocástico presenta una mejora cercana al 13% en relación a los beneficios del determinístico, para ambos escenarios. Se concluye que de todas formas hay un aumento del beneficio, de forma que solo queda estudiar hasta qué punto es útil seguir aumentando el número de escenarios a incluir en el modelo.

7.3 Comparación entre Formulación Extendida y Progressive Hedging

Se contrastaron los resultados de la implementación extendida del problema y la mejor configuración de PH. Para las instancias pequeñas, menores a los 100 escenarios, ambas formulaciones tienen un comportamiento similar, tanto en tiempo como en el Gap alcanzado. En las instancias menos masivas, hasta los 300 escenarios, PH no pudo superar la formulación extendida. Esto es esperable dado que para pocos datos la resolución por iteraciones del algoritmo toma más tiempo que resolver todo de forma simultánea.

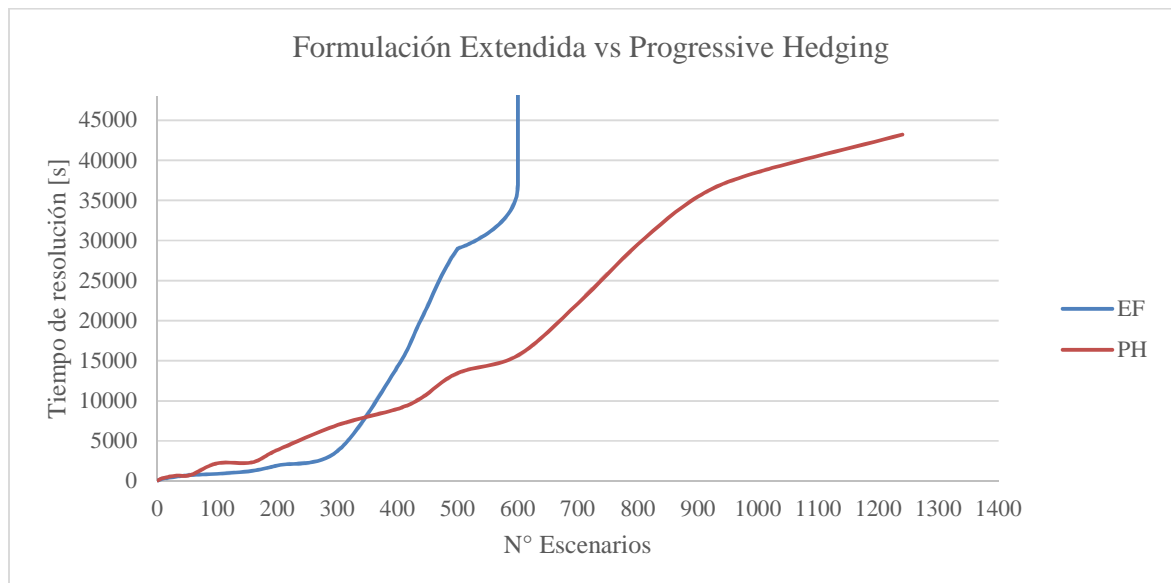
Por otra parte, se encuentra el número de escenarios para el cual la versión extendida no puede ser resuelta con los recursos computacionales disponibles, al punto que ni siquiera es posible realizar la carga de las instancias, dado el gran tamaño de las mismas. Este punto es entre 600 y 700 escenarios. En la Figura 7-3 se señalan los resultados obtenidos, con los tiempos de resolución y el Gap correspondiente. El Gap calculado para PH proviene de una estimación realizada entre el valor objetivo y el Gap entregado por CPLEX.

Figura 7-3 Comparación entre Formulación Extendida y Progressive Hedging

Escenarios	Extendido		PH	
	Tiempo [s]	Gap [%]	Tiempo [s]	Gap [%]
1	74	0,01%	74	0,01%
10	301	0,01%	389	0,01%
32	547	0,02%	684	0,02%
55	749	0,01%	751	0,02%
100	905	0,24%	2217	0,30%
158	1270	0,14%	2335	0,054%
200	1947	0,23%	3879	0,65%
300	3755	0,54%	6987	0,80%
400	14289	0,87%	9005	0,80%
440	20302	1,19%	10388	1,20%
500	28996	1,40%	13478	1,44%
600	36875	1,79%	15670	1,78%
700	Sin solución	-	22157	1,94%
800	Sin solución	-	29545	1,89%
900	Sin solución	-	35488	2,01%
1000	Sin solución	-	38541	2,54%
1240	Sin solución	-	43200	3,12%

En la Figura 7-4 se grafican los resultados para todas las instancias. En el mismo se puede apreciar en qué punto la formulación extendida no puede entregar resultados, mientras que el algoritmo PH sigue dando soluciones. Más aún, se estabiliza en cuanto a la pendiente de los tiempos de resolución a partir de los 800 escenarios. Esto se entiende porque el problema se vuelve más complejo ya que deben leerse más datos, pero la solución no cambia: el espectro de climas que cubre un árbol de 800 y uno de 1000 escenarios no varía mucho, solo varía la cantidad de datos que deben contabilizarse.

Figura 7-4 Gráfico para los resultados de la comparación



Toma mucha relevancia la búsqueda de opciones para resolver el algoritmo PH para instancias mayores a los 600 escenarios, dado que si bien es posible encontrar soluciones, el costo en relación al tiempo que toma encontrarlas es demasiado alto y sigue aumentando al incrementar el tamaño de las instancias. En las siguientes secciones se muestran las alternativas de clustering y bundling.

7.4 Caso clustering de escenarios

7.4.1 Conglomerados de K medias

Se implementó el algoritmo de conglomerados de K medias, en primer lugar para la selección de los escenarios representativos de la distribución y luego se mantuvo el 5% de los extremos para evitar que se excluyeran casos que pueden resultar infactibles si no son considerados. Con esto, se modifica la instancia para luego utilizar la mejor configuración para PH utilizada en los experimentos anteriores.

Para cada instancia se hizo una reducción al 25% y al 50% de la original, comparando los mismos indicadores de tiempo y diferencia con la función objetivo original (para la instancia de 10 escenarios no tenía sentido reducir al 25% ya que resultaba una instancia demasiado reducida para hacer cualquier análisis). Los resultados se presentan de forma que para cada instancia están las dos instancias reducidas, el tiempo que toma obtener una solución, y qué proporción representa eso respecto al tiempo que toma resolver la instancia original, y por último la función objetivo, presentada como la variación con respecto al valor para la instancia original

Para las instancias menos masivas, la reducción no resulta demasiado útil, como se muestra en los resultados en la Figura 7-5. Si bien la reducción de tiempo es importante en muchos casos, la variación del valor de la función objetivo resulta bastante significativa.

Figura 7-5 Resultados de Clustering con conglomerados de K medias para instancias pequeñas

Escenarios	Instancia reducida	Tiempo reducido [s]	Tiempo reducido [%]	Diferencia F. Objetivo [%]
10	5	177	45,45%	5,29%
	32	385	56,24%	3,24%
32	8	314	45,87%	3,89%
	55	399	53,11%	2,65%
55	14	299	39,87%	3,05%
	100	1147	51,74%	3,58%
100	25	911	41,10%	4,51%
	158	1304	55,83%	2,22%
158	40	939	40,22%	2,57%
	200	1284	33,11%	2,49%
200	50	455	11,74%	2,21%
	300	4978	71,24%	2,33%
300	75	3486	49,89%	2,98%

Por otra parte, la reducción en instancias pequeñas resulta menos útil en comparación a las instancias masivas, como se muestra en la Figura 7-6. Para estos casos, la reducción de tiempo puede llegar al 40%, pero variando menos de un 1% de la función objetivo. Esto era esperable, ya que reducir un árbol de escenarios que ya es pequeño y por lo tanto cubre un espectro reducido de climas, acotarlo aún más ve mermado el resultado en la calidad de la solución. Mientras que una instancia masiva puede ser reducida y seguir representando la misma distribución anterior.

Figura 7-6 Resultados de Clustering con conglomerados de K medias para instancias masivas

Escenarios	Instancia reducida	Tiempo reducido [s]	Tiempo reducido [%]	Diferencia F. Objetivo [%]
400	200	3795	42,14%	1,06%
	100	855	9,50%	3,65%
440	220	4338	41,76%	1,03%
	110	808	7,78%	3,62%
500	250	8083	59,97%	1,11%
	125	743	5,51%	3,14%
600	300	8498	54,23%	1,25%
	150	505	3,22%	3,24%
700	350	9955	44,93%	1,16%
	175	707	3,19%	3,27%
800	400	12069	40,85%	0,89%
	200	810	2,74%	4,01%
900	450	15168	42,74%	0,58%
	225	405	1,14%	3,98%
1000	500	20519	53,24%	0,24%
	250	378	0,98%	3,17%
1240	620	18338	42,45%	0,21%
	310	441	1,02%	2,14%

7.4.2 Clusters de tamaño uniforme

Se selecciona de manera uniforme los escenarios como se indica en la metodología explicada en la subsección 5.2.7. De igual forma que en para el algoritmo modificado de los conglomerados de K medias, se mantienen los extremos del árbol de escenarios y se hacen reducciones al 25% y al 50% para cada instancia.

El gran beneficio de esta técnica por sobre la anterior, es la selección de los escenarios. Es bastante más simple y por lo tanto más rápida de llevar a cabo.

En cuanto a los tiempos de resolución del modelo y las diferencias con la función objetivo, los resultados varían levemente, y no lo suficiente para concluir si un método resulta mejor que otro. Los resultados para esta metodología se muestran en la Figura 7-7.

Figura 7-7 Resultados de clustering de tamaño uniforme

Escenarios	Instancia reducida	Tiempo reducido [s]	Tiempo reducido[%]	Diferencia F. Objetivo[%]
10	5	184	47,40%	5,29%
32	16	375	54,84%	3,24%
	8	307	44,98%	3,89%
55	28	399	53,10%	2,65%
	14	301	40,10%	3,05%
100	50	1118	50,41%	3,50%
	25	831	37,50%	4,12%
158	79	1222	52,34%	2,40%
	40	898	38,45%	2,13%
200	100	1323	34,10%	2,49%
	50	465	12,00%	2,21%
300	150	5062	72,45%	2,24%
	75	3493	49,99%	2,16%
400	200	4042	44,89%	2,16%
	100	1027	11,40%	3,89%
440	220	4702	45,26%	1,45%
	110	716	6,89%	3,89%
500	250	6719	49,85%	1,06%
	125	659	4,89%	3,47%
600	300	8173	52,16%	1,14%
	150	486	3,10%	3,01%
700	350	9297	41,96%	1,14%
	175	667	3,01%	3,20%
800	400	11688	39,56%	1,24%
	200	948	3,21%	5,48%
900	450	15561	43,85%	0,58%
	225	422	1,19%	3,98%
1000	500	21386	55,49%	1,25%
	250	567	1,47%	3,89%
1240	620	21427	49,60%	1,24%
	310	816	1,89%	2,25%

7.5 Caso bundling de escenarios

A partir de lo señalado en la subsección 5.2.8, se realiza una serie de experimentos basados en la creación de bundles de escenarios. Se utilizaron 2 técnicas: la creación de bundles aleatorios y por otra parte a partir de la utilización del algoritmo de los conglomerados de K medias, de forma de identificar los escenarios que presentan características similares y por lo tanto serían más aptos para ser trabajados en conjunto.

El caso de la creación de bundles aleatorios se implementa directamente, ya que está incluido en el paquete Coopr, por lo mismo resulta sencillo hacer los experimentos, solo es necesario ingresar el número de agrupaciones que desean construirse. Para las instancias de prueba, se observa que los tiempos de resolución mejoran en la medida que se aumenta la cantidad de bundles a generar, pero no alcanzan a mejorar los resultados. En particular, los mejores resultados

en relación al tiempo se obtienen al construir un número de grupos bastante similar a la cantidad de escenarios de la instancia original, lo que guarda relación con que los “macro-escenarios” resultan más complejos de resolver que los escenarios originales. En conclusión, no vale la pena el tiempo extra que toma la inicialización de esta metodología, en comparación a las mejoras poco significativas en el tiempo.

Por otra parte, al utilizar el algoritmo de los conglomerados, toma un tiempo de procesamiento el generar el archivo que define los grupos, archivo que debe ser leído por el programa para empezar la inicialización. En este caso se podrían esperar mejores resultados, dado que dentro de un mismo grupo los escenarios compartirían características y debieran ser más sencillos para el solver. El problema que se presentó con esta técnica recae en la inicialización, esto es la creación de los macro-escenarios. A modo de experimentación, no vale la pena intentar esta metodología, ya que llegó a tomar hasta 5 horas solamente en la etapa previa a la resolución. No se puede concluir en relación a la mejor técnica para agrupar los escenarios, dado que ni siquiera se llegó a obtener una solución.

7.6 Factibilidad y Calidad de la Solución

Según la implementación ya expresada en la sección 6.3 se obtienen los resultados de factibilidad para cada una de las soluciones encontradas. La calidad de cada árbol se definirá entonces como el porcentaje de error, y ese error corresponde a la cantidad de escenarios infactibles.

En la Figura 7-8 se encuentra el factor de infactibilidad para cada una de las instancias trabajadas. Es claro que a mayor cantidad de escenarios, mejor es la representatividad del universo total disponible, y por lo tanto la probabilidad de encontrar escenarios disponibles disminuye considerablemente.

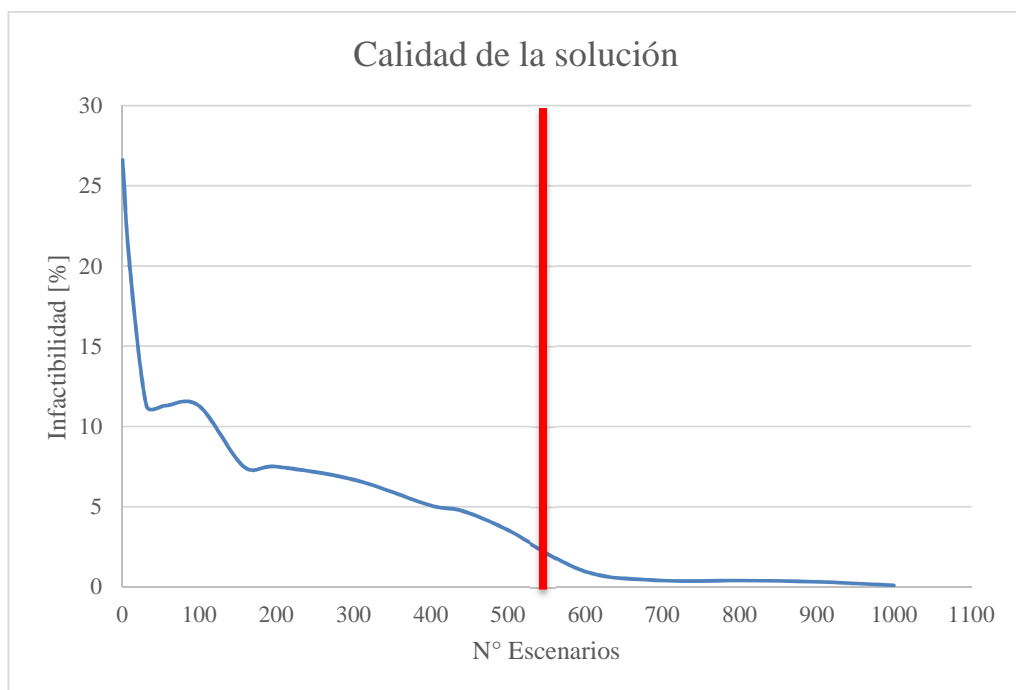
Figura 7-8 Nivel de infactibilidad por instancia

N° Escenarios	Factor de infactibilidad [%]
1	26,61%
10	20,16%
32	11,29%
55	11,29%
100	11,29%
158	7,5%
200	7,5%
300	6,69%
400	5,08%
440	4,76%
500	3,55%
600	0,97%
700	0,4%
800	0,4%
900	0,32%
1000	0,1%

En este caso se ve claramente cómo ayuda la utilización de un modelo estocástico por sobre uno determinístico (que sería el caso de un solo escenario) para obtener beneficios en cuanto a la solución encontrada. Para más de un 26% de los escenarios totales, la solución determinística ni siquiera resulta factible, adicionalmente a que la solución es óptima solo para dicho caso, el resto de escenarios factibles igualmente pueden no resultar en el aprovechamiento de los recursos forestales de la mejor forma que se pueda.

En la Figura 7-9 se grafican los resultados anteriores, y se observa claramente en qué punto el factor de infactibilidad se estabiliza. Se puede concluir que dado que para cierto nivel la cantidad de escenarios infactibles no continua incrementándose de forma considerable, no sería necesario dedicar mayor cantidad de recursos para resolver más escenarios si la calidad de la solución no mejorará. Este punto se encuentra alrededor de los 650 escenarios.

Figura 7-9 Gráfico para el factor de infactibilidad



Así, se puede decir que un árbol de 650 escenarios sería suficiente para estar cubierto frente a la incertidumbre, pero aprovechando de buena forma los recursos computacionales que están a disposición del público en general y en un tiempo razonable, utilizando las mejoras propuestas para el algoritmo Progressive Hedging.

7.6.1 Factibilidad para el caso de Clustering

Considerando que puede comprobarse la utilidad de usar la técnica de clustering para resolver el problema forestal, es importante concluir también sobre la calidad de la solución obtenida con dicho método. Así, se realiza el mismo procedimiento, pero esta vez viendo qué porcentaje de la instancia original para cada experimento con clustering resulta factible dada la solución encontrada para la instancia reducida.

Esto quiere decir, por ejemplo, en la primera instancia de 10 escenarios y su correspondiente reducción a 5 escenarios, qué parte de los 10 escenarios es factible dada la solución encontrada para la instancia reducida.

De este análisis, los resultados se muestran en la Figura 7-10. Similar a lo que se concluyó en relación a los resultados generales y la comparación con la instancia de 1240 escenarios, mientras más grande la instancia que se trabaja, mejora la factibilidad, ya que la representatividad del espectro de escenarios sigue siendo alta.

Figura 7-10 Factibilidad para instancias reducidas por clustering

Escenarios	Instancia reducida	Infactibilidad para K medias	Infactibilidad para uniforme
10	5	30,00%	30,00%
32	16	21,88%	21,88%
	8	25,00%	25,00%
55	28	18,18%	18,18%
	14	21,82%	21,82%
100	50	14,00%	13,00%
	25	19,00%	18,00%
158	79	9,49%	8,86%
	40	13,92%	13,29%
200	100	7,00%	7,00%
	50	14,50%	14,50%
300	150	5,33%	4,67%
	75	8,00%	7,33%
400	200	4,75%	5,00%
	100	6,75%	7,00%
440	220	4,09%	4,77%
	110	6,14%	6,82%
500	250	3,40%	2,80%
	125	5,20%	4,60%
600	300	3,17%	3,00%
	150	4,67%	4,50%
700	350	2,57%	2,14%
	175	3,57%	3,29%
800	400	1,75%	1,63%
	200	2,63%	2,50%
900	450	1,44%	1,44%
	225	1,67%	1,67%
1000	500	0,90%	1,10%
	250	1,30%	1,50%
1240	620	0,40%	0,48%
	310	1,21%	1,29%

En la Figura 7-11 se grafica la evolución del porcentaje de infactibilidad para la técnica de conglomerados de K medias, a medida que aumenta el número de escenarios, tanto para las reducciones a 25% como para 50%. En dicho gráfico queda claro que al aumentar el tamaño de la instancia original, el error disminuye drásticamente, tanto para las reducciones de 50% como 25%. Por otra parte, también se nota que la diferencia para ambas reducciones disminuye de forma considerable a medida que aumenta el número de escenarios. Resultados similares se observan para la segunda técnica de clustering, mostrados en la Figura 7-12.

Figura 7-11 Gráfico de infactibilidad para caso Clustering con K medias

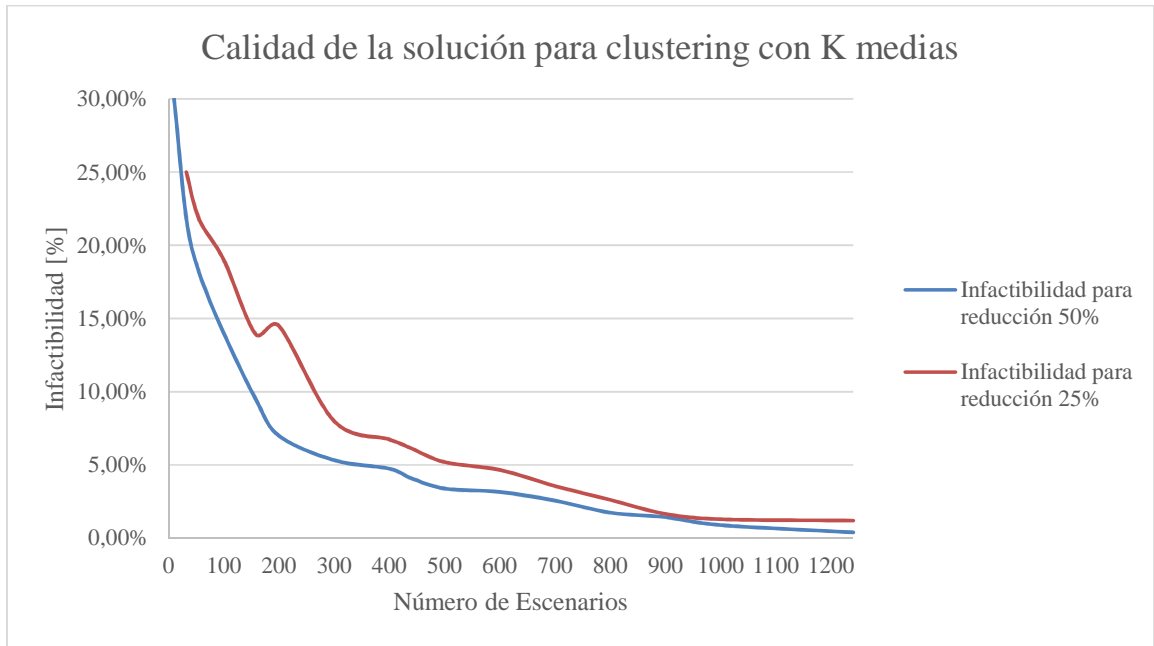
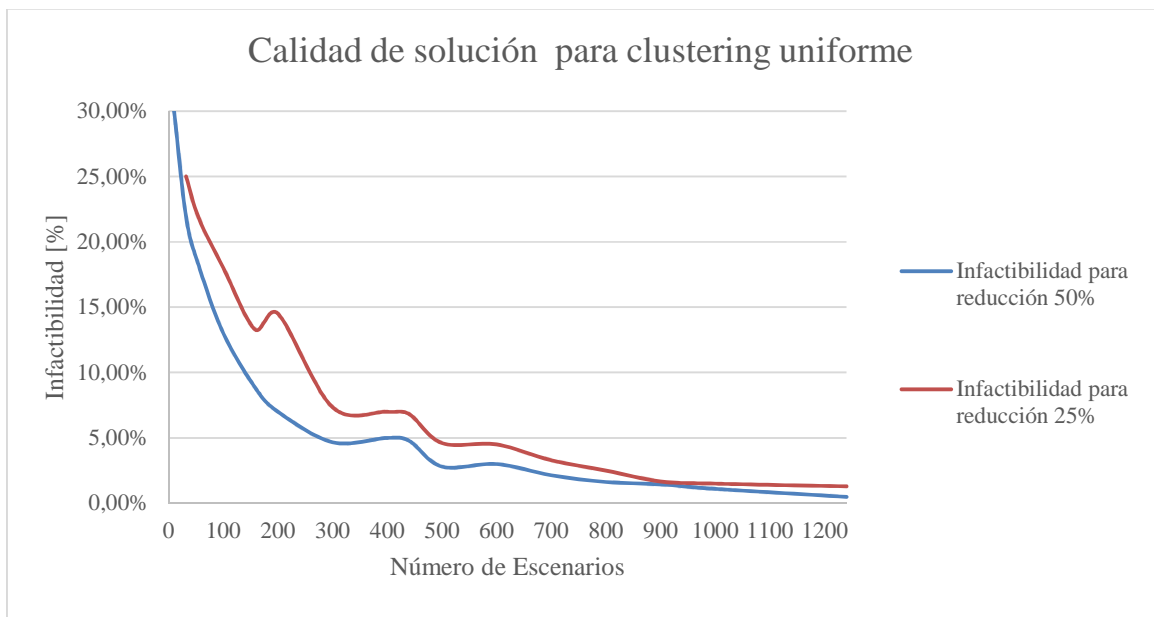


Figura 7-12 Gráfico de infactibilidad para caso Clustering uniforme



A partir de esto último se puede concluir que al momento de decidir si realizar una reducción de la instancia original con la técnica de clustering, se debe tomar en cuenta que a pesar de que el error en la función objetivo es bastante mayor para la reducción a 25%, el tiempo también disminuye considerablemente y la factibilidad no varía demasiado en relación a una reducción de 50%. Por lo tanto, tomar una muestra pequeña de la instancia original puede resultar provechoso, sin mermar demasiado el resultado.

Capítulo 8

Conclusiones

El desarrollo de este trabajo de memoria incluye la implementación de un modelo forestal que representa un bosque real en la zona central de Portugal, incluyendo el factor climático como variable estocástica. Su importancia radica en la relevancia que ha tenido en los últimos años el cambio climático que sufre el planeta y que por lo tanto afecta las decisiones en este ámbito, es importante considerar planes de acción para un rango mucho más amplio de posibles escenarios de clima.

La complejidad aumenta en relación a trabajos anteriores, incrementando considerablemente la cantidad de escenarios a estudiar, y adicionalmente, los escenarios generados están sustentados por una base de predicciones reales de clima para distintas zonas de Portugal, por lo que deben representar de buena manera el clima esperado para los próximos años. La metodología propuesta puede ser replicada para otras zonas del mundo, teniendo en cuenta siempre que será necesario contar con predicciones realizadas por centros climatológicos, con climatólogos expertos.

Se demuestra la utilidad del algoritmo Progressive Hedging a la hora de resolver problemas de carácter estocástico y con múltiples etapas, con gran cantidad de variables y restricciones, tomando instancias de gran magnitud. Se incluye el detalle de las mejoras y ajustes que se pueden hacer al algoritmo para aumentar su rendimiento, principalmente en términos de tiempos de resolución.

La mejor configuración encontrada para el algoritmo PH es con un punto de corte de $\varepsilon = 0,01$, con un máximo de 5 iteraciones. Se fijan variables luego de la primera iteración, el valor para el factor de penalización es $\rho = 0,0001$, linealización del término proximal en 2 tramos y se paraleliza el algoritmo en 4 núcleos.

Las instancias testeadas fueron resueltas satisfactoriamente por el algoritmo, aumentando de 32 escenarios, en el trabajo que antecede a esta memoria, a 1240 escenarios climáticos, para el mismo bosque en estudio.

La formulación extendida en CPLEX logra ser superior solo para instancias menores a los 300 escenarios, precisamente porque en esos casos el problema es menos masivo y por lo tanto la metodología del algoritmo PH no logra superarlo. Sin embargo, al tratar con instancias mayores, la formulación extendida ni siquiera es capaz de llegar a un resultado, esto para una cantidad mayor a 600 escenarios.

La técnica de clustering de escenarios permite reducir la dimensionalidad del problema, esto tomando instancias masivas para llevarlas a tamaños manejables, pero que a su vez mantenga el carácter estocástico del problema original y por lo tanto que sus resultados sean robustos. Tanto el uso de conglomerados de K medias como el escoger escenarios de manera uniforme son fácilmente implementados, y además presentan buenos resultados en cuanto a los tiempos de resolución. En particular la segunda técnica enunciada es más sencilla en cuanto a la selección de los escenarios, por lo que presentaría una ventaja respecto a la primera. Dado que los tiempos al implementar PH, incluso con las mejoras planteadas, puede tomar mucho tiempo, es muy positivo tener técnicas complementarias.

La técnica de bundling no representó una mejora para las instancias masivas. Por un lado, la técnica de crear grupos aleatorios no presentó una mejora significativa en el tiempo de resolución, y por otra parte no fue posible hacer los experimentos para la creación de grupos basado en el algoritmo de conglomerados de K medias, ya que solo el tiempo de inicialización superó el tiempo de todas las otras alternativas.

Se logra dar respuesta para el modelo en estudio a un gran desafío, como lo es encontrar el número óptimo de escenarios con los cuales trabajar. Esto basado en sustentos matemáticos, y que permita equiparar el aumento significativo en los tiempos de resolución con la calidad de la solución. Se determina que el número eficiente se encuentra entre los 600 y los 700 escenarios, ya que en este punto no vale la pena tratar de resolver instancias mayores ya que los niveles de infactibilidad no disminuyen.

Teniendo en cuenta que alrededor de los 600 escenarios se encuentra el punto de estabilidad de infactibilidad, se puede analizar la posibilidad de usar la técnica de clustering, que resultó efectiva en gran parte de las instancias masivas. Así, se podría pasar de tiempo de resolución de 4 horas 20 minutos, a 2 horas 20 minutos si se reduce la instancia a la mitad, e incluso se puede llegar a un tiempo de 8 minutos, teniendo un error de 3,24% en la función objetivo, al reducir la instancia de 600 a 150 escenarios con la técnica de conglomerados de K medias.

Finalmente, se concluye que se cumple con el objetivo de maximizar el valor esperado de la función objetivo para el plan de producción, implementando correctamente el algoritmo previsto y permitiendo una buena respuesta frente a instancias de gran tamaño. Particularmente, se presentan grandes beneficios en relación a la inclusión de un amplio rango de climas en estudio, en comparación a trabajos anteriores y avances en cuanto al análisis del número de escenarios que será significativo.

8.1 Trabajo futuro

En primer lugar, la inclusión de otras fuentes de incertidumbre permitirá generar un modelo cada vez más completo y robusto frente a la realidad que enfrenta el tomador de decisiones. Entre dichas fuentes se puede encontrar la probabilidad de incendios forestales, y por lo tanto la decisión de incluir cortafuegos, o diferentes estacionalidades, en tal caso sería necesario hacer un plan mensual que tome en consideración la variación de las condiciones climáticas durante el año. Además, se dejó de lado la variabilidad inherente al crecimiento de bosques, para un mismo clima sí existe incertidumbre en el crecimiento, factor que se asumió como conocido en las simulaciones realizadas.

En la formulación del problema también pueden incluirse otros factores, como la inclusión de los caminos que también implican tala de árboles, y la adyacencia. Ésta última hace referencia a restricciones ambientales, donde se debe tomar en consideración la cercanía entre las áreas a talar en cada periodo y la restricción que evite que existan zonas demasiado amplias que sean cosechadas simultáneamente.

Por otro lado, hay otras posibilidades en relación a los parámetros del problema, se pueden incluir distintos tipos de demanda, o por cierto la posibilidad de mantener inventario de un periodo a otro. Un factor que no ha sido tomado en cuenta, son los distintos costos involucrados en la producción de la madera, esto podría afectar directamente los beneficios, por ejemplo, al tomar en cuenta la mano de obra, la maquinaria a utilizar, entre otros.

También hay trabajo pendiente en cuanto a la implementación de Progressive Hedging. La implementación de dicho algoritmo presenta grandes oportunidades, en particular al paralelizar la resolución en distintos procesadores o núcleos. En esta memoria solo se utilizó una máquina para resolver cada problema. Esto conlleva problemas, ya que se depende de la memoria RAM de cada computador, y utilizar más núcleos en un mismo equipo puede llevar a que la memoria se agote y pierda eficiencia. Mientras que si, incluso usando la misma cantidad de núcleos pero en distintas máquinas, se podría mejorar el rendimiento del algoritmo.

Además, se esperaría poder realizar avances en la investigación de técnicas para mejorar la comunicación entre los núcleos y así disminuir sustancialmente el tiempo de cómputo. Ya que si bien puede ser útil resolver el problema en distintos computadores, en ciertos casos, en particular de instancias mayores, se presentan demoras al comunicar los distintos procesadores y luego reunir los resultados en el computador maestro.

Finalmente, se presenta como desafío seguir el estudio del número óptimo de escenarios a utilizar y expandir la metodología para encontrar este número a otro tipo de problemas, más allá de los problemas forestales. Queda pendiente también seguir comprobando la efectividad de usar un amplio rango de escenarios y su utilidad práctica, teniendo por ejemplo, la posibilidad de implementar alguno de los planes de producción propuestos y concluir de manera empírica qué beneficios conlleva.

Capítulo 9

Glosario

- GAP o MIPGAP: Distancia en términos porcentuales existente entre la mejor solución factible calculada y la mejor cota del óptimo que ha sido encontrada por el solver.
- MIP: Mixed Integer Programming. Problema de optimización lineal que hace referencia a un modelo que incluye variables y enteras (incluida la presencia de variables binarias).
- EF: Extended Formulation. Hace referencia a la formulación extendida, incluye todos los escenarios y resueltos de forma simultánea.
- Formulación Extendida Explícita: Incluye todos los escenarios, con variables por escenario y restricciones de no anticipatividad explícitas.
- Formulación Extendida Compacta: Incluye todos los escenarios, con variables por nodo de información, lo que implementa implícitamente la no anticipatividad.
- PH: Progressive Hedging. Hace referencia al algoritmo de descomposición por escenarios utilizado en esta memoria. En particular, la traducción correcta al español dado el contexto en que se utiliza, corresponde a “protegerse del riesgo”.
- SOLVER: Corresponde a una rutina de software que se encarga de resolver un problema en particular. Dentro del contexto de optimización, hacer referencia a un software que utiliza distintos algoritmos para resolver los problemas de optimización.
- ForEAdapt: Knowledge Exchange between Europe and America on forest growth models and optimization for adaptive forestry. Programa en el cual está enmarcado este trabajo de memoria, participan universidades de Chile, Brasil, Portugal, Finlandia, España, Suecia y Estados Unidos.

Capítulo 10

Bibliografía

- [1] “IFN6 – Áreas de uso de suelo de las especies forestales de Portugal continental. Resultados preliminares”. [pdf], 43 pp, Instituto de Conservación de la naturaleza y los bosques. Lisboa, 2013.
- [2] “ForEAdapt: Knowledge Exchange between Europe and America on forest growth models and optimization for adaptive forestry” [Online] Disponible: <http://www.isa.utl.pt/cef/pub/ForEAdapt/>
- [3] J. Garcia-Gonzalo, A. Weintraub. “A Stochastic Optimization Model to Adressing Climate Change in Forest Planning” presentado en 20º Conferencia de la Federación Internacional Sociedades de Investigación Operacional, Barcelona, España, 2014.
- [4] Sandia National Laboratories, software Package Coop <https://software.sandia.gov/trac/coopr>
- [5] C. Ortiz, S. Varas, J. Vera. Optimización y Modelos para la Gestión. Dolmen, 2001.
- [6] R.T. Rockafellar, R. J-B. Wets. “Scenario and policy aggregation in optimization under uncertainty”, Mathematics of Operations Research, 16, pp 119-147, 1991.
- [7] H. Heitsch, W. Römisch. “Scenario Reduction Algorithms in Stochastic Programming”, Computational Optimization and Applications, 24, pp 187-206, 2003.
- [8] J.P. Watson, W. Hart, D.L. Woodruff “PySP: Modeling and Solving Stochastic Programs in Python”, a aparecer en Mathematical Programming Computation, 2011.
- [9] ENSEMBLES project <http://www.ensembles-eu.org/>
- [10] M. Quinteros, A. Alonso, L. Escudero, M. Guignard, A. Weintraub “Forestry Management under Uncertainty” Annals of Operations Research, 1572-9888, 2009.
- [11] C. Pais “Complete stochastic forestry planning problem using Progressive Hedging Algorithm” Tesis, Universidad de Chile, Santiago, Chile, 2014.
- [12] Python Pickle Package Documentation <https://doc.python.org/2.7/library/pickle.html>
- [13] F. Badilla “Problema de Planificación Forestal Estocástico resuelto a través del algoritmo Progressive Hedging” Tesis, Universidad de Chile, Santiago, Chile, 2010.

Capítulo 11

Anexos

Anexo A. Teorema de Redistribución de probabilidades

Para determinar la forma en la que se deben distribuir las probabilidades de ocurrencia de los escenarios no seleccionados, se utiliza el siguiente teorema, de forma de mantener la distribución probabilística original. De manera simple, se agrega la probabilidad de ocurrencia a aquellos escenarios que se encuentren más cercanos.

Dado un conjunto $J \subset \{1, \dots, N\}$ se tiene el siguiente problema de optimización:

$$D_j = \min \left\{ D(J; q) : q_j \geq 0, \sum_{j \notin J} q_j = 1 \right\} = \sum_{t \in J} p_t \min_{j \notin J} c(w_t, w_j)$$

Entonces, el óptimo se alcanza en:

$$\bar{q}_j = p_j + \sum_{t \in J_j} p_t \quad \forall j \notin J$$

$$J_j := \{t \in J : j = j(t) \in \operatorname{argmin}_{j \notin J} c(w_t, w_j)\} \quad \forall t \in J$$

Anexo B. Extracto de los datos entregados por el programa ForEAdapt

La información que incluye la base de datos define el clima de cada estación, para cada año, para cada mes. Se escoge trabajar con la precipitación (variable “Rain”) ya que para un mismo nivel de precipitación, el resto de las variables se mantiene relativamente constante. En la Figura 11-1 se muestra un extracto para una estación meteorológica de Portugal, entre los años 2014 y 2018.

Figura 11-1 Extracto de los datos para una estación meteorológica

Station	Year	Month	Tmean	Tmax	Tmin	Rain	SolarRad	RainDays	FrostDays	VPD	Evap	RelHum
448	2014	1	15,3	16,2	14,3	51,7	8,6	11	0	999	2,3	81,5
448	2014	2	14,4	15,7	13,1	16,4	13,3	5	0	999	4,2	72,9
448	2014	3	16,1	17,3	14,9	7,8	17,8	1	0	999	3,1	77,9
448	2014	4	16,7	18	15,3	2,6	24,9	0	0	999	2,8	75,4
448	2014	5	18,1	19,3	16,9	5,9	27,5	2	0	999	3,4	77,9
448	2014	6	19,1	20,6	17,7	1,2	28,9	1	0	999	3,6	74,5
448	2014	7	21,8	23,5	20,2	8,3	26	3	0	999	1,2	82,6
448	2014	8	22,3	24,1	20,6	2,5	24,2	1	0	999	1,5	81
448	2014	9	21,5	22,8	20,3	0,5	20,8	0	0	999	1,9	81,5
448	2014	10	19,7	21,1	18,3	3,4	14,3	0	0	999	2,8	78,8

448	2014	11	18,2	19,1	17,2	44,6	10	7	0	999	3,1	76,9
448	2014	12	17,1	18,1	16	45,7	7,4	9	0	999	2,9	77,9
448	2015	1	15,6	16,5	14,6	193,8	7,5	18	0	999	3,9	76,3
448	2015	2	15,3	16,4	14,1	20,2	12,5	3	0	999	3,7	79,1
448	2015	3	15,3	16,3	14,3	116,9	15,6	11	0	999	3,3	78,3
448	2015	4	17	18,6	15,4	2,1	24,3	0	0	999	3,3	74,2
448	2015	5	18,5	19,7	17,4	17,2	24,4	4	0	999	3,1	78,6
448	2015	6	20,5	22,1	18,8	3,8	27,2	2	0	999	3,3	77,4
448	2015	7	22,4	24	20,7	0,8	28	0	0	999	2,6	77,3
448	2015	8	22,7	24,1	21,4	4,4	25	1	0	999	3,3	77,7
448	2015	9	22	23,6	20,5	2,8	19,7	0	0	999	4,1	76,9
448	2015	10	20,6	21,9	19,4	11,3	14	2	0	999	2,8	81,3
448	2015	11	17,5	18,6	16,4	88,3	9,8	9	0	999	4,7	75
448	2015	12	14,9	16,1	13,7	35,6	8,7	4	0	999	4,5	77
448	2016	1	16,1	17	15,1	36	8,4	6	0	999	1,8	83,6
448	2016	2	15,2	16,7	13,8	1	14,1	0	0	999	2,8	79,5
448	2016	3	14,9	16,6	13,2	2,1	18,5	0	0	999	4,3	72,1
448	2016	4	17,7	19,2	16,1	1,2	24,9	0	0	999	2,2	77,6
448	2016	5	19,2	21,1	17,3	0,9	27,6	0	0	999	3,6	73,2
448	2016	6	21,7	23,1	20,4	1,9	29,3	0	0	999	2,2	81
448	2016	7	22,4	23,8	21	15,8	27,5	2	0	999	2,7	78,4
448	2016	8	22,9	24,3	21,5	5	22,9	2	0	999	1,3	84,3
448	2016	9	23,2	24,6	21,9	2,8	20,5	1	0	999	2,8	78
448	2016	10	20,6	21,9	19,4	35,5	15,2	6	0	999	4,4	74,1
448	2016	11	18,6	19,8	17,5	166,5	9,4	13	0	999	5	76,1
448	2016	12	16,7	17,9	15,6	74,9	7,4	8	0	999	4,6	76,3
448	2017	1	16,3	17,3	15,2	35,7	8,5	5	0	999	3	80,7
448	2017	2	15	16,4	13,6	12,3	14,2	4	0	999	4,6	74,5
448	2017	3	16,2	17,6	14,8	7,1	17,8	4	0	999	4,3	72
448	2017	4	15	16,7	13,3	23,8	23,3	6	0	999	5,4	68,4
448	2017	5	19	20,2	17,8	2,8	27	0	0	999	2,8	79,6
448	2017	6	22,2	23,9	20,4	0,6	28,9	0	0	999	3,5	73,5
448	2017	7	23,6	25,2	22,1	1,3	28,3	0	0	999	2,9	74,2
448	2017	8	23	24,5	21,5	0,9	25	0	0	999	2,1	78,6
448	2017	9	22,4	23,8	21,1	5,5	20,4	2	0	999	2,5	79,2
448	2017	10	21	22,2	19,8	14,1	15	3	0	999	3,4	77,3
448	2017	11	18,6	19,8	17,3	31,8	10,8	5	0	999	3,5	75,5
448	2017	12	17,1	18	16,1	0,9	8,9	0	0	999	3,9	77,4
448	2018	1	14,9	16,3	13,4	29,4	9,6	7	0	999	4,1	73,6
448	2018	2	15,1	16,2	14	12,2	13	6	0	999	3,2	77,5
448	2018	3	15,4	16,8	14	3,3	18,5	1	0	999	4,2	74,4
448	2018	4	17,3	18,9	15,7	5,6	24,5	2	0	999	3,1	72,4
448	2018	5	17,5	18,9	16	2,7	28,2	1	0	999	4	74,2

448	2018	6	20,6	21,9	19,2	1	28,5	0	0	999	2	79,9
448	2018	7	21,9	23,2	20,6	1	27,7	0	0	999	1,8	81,2
448	2018	8	22,6	24,4	20,8	1,2	25,1	0	0	999	2	78,4
448	2018	9	22,5	23,6	21,4	13,3	19,3	4	0	999	1,7	83,5
448	2018	10	20,4	22	18,8	10,6	15,5	1	0	999	4,2	73
448	2018	11	18,8	19,9	17,8	49,1	9,7	6	0	999	3,3	79,9
448	2018	12	17,1	18,2	15,9	29,2	8,1	4	0	999	3,5	78,8

Anexo C. Resultados para distintos niveles de demanda para todas las instancias

Figura 11-2 Resultados para distintos niveles de demanda para todas las instancias

Escenarios	Demanda	Formulación Extendida		Progressive Hedging	
		Tiempo[s]	Gap [%]	Tiempo[s]	Gap [%]
1	0	60,1	0,01%	60,1	0,01%
	10000	61,2	0,01%	61,2	0,01%
	20000	54,7	0,01%	54,7	0,01%
	30000	59,7	0,01%	59,7	0,01%
	40000	55	0,01%	55	0,01%
	50000	53,2	0,01%	53,2	0,01%
	60000	59,4	0,01%	59,4	0,01%
	70000	67,2	0,01%	67,2	0,01%
	80000	74,2	0,01%	74,2	0,01%
10	0	154	0,01%	204	0,01%
	10000	159	0,01%	209	0,01%
	20000	201	0,01%	215	0,01%
	30000	213	0,01%	236	0,01%
	40000	224	0,01%	227	0,01%
	50000	251	0,01%	289	0,01%
	60000	274	0,01%	324	0,01%
	70000	268	0,01%	333	0,01%
	80000	301	0,01%	389	0,01%
32	0	334	0,02%	451	0,02%
	10000	374	0,02%	444	0,02%
	20000	385	0,02%	498	0,02%
	30000	420	0,02%	512	0,02%
	40000	489	0,02%	533	0,02%
	50000	501	0,02%	547	0,02%
	60000	532	0,02%	598	0,02%
	70000	549	0,02%	625	0,02%
	80000	547	0,02%	684	0,02%

		Formulación Extendida		Progressive Hedging	
Escenarios	Demanda	Tiempo[s]	Gap [%]	Tiempo[s]	Gap [%]
55	0	622	0,01%	675	0,01%
	10000	624	0,01%	661	0,01%
	20000	635	0,01%	669	0,01%
	30000	641	0,02%	645	0,02%
	40000	623	0,01%	639	0,01%
	50000	645	0,01%	689	0,01%
	60000	655	0,01%	656	0,01%
	70000	644	0,01%	673	0,01%
	80000	749	0,01%	751	0,02%
100	0	701	0,11%	874,2	0,10%
	10000	722	0,15%	836	0,21%
	20000	731	0,15%	957	0,18%
	30000	754	0,20%	960	0,24%
	40000	789	0,21%	1001	0,26%
	50000	821	0,17%	1481	0,19%
	60000	835	0,14%	1244	0,16%
	70000	866	0,23%	2178	0,29%
	80000	905	0,24%	2217,3	0,30%
158	0	1142	0,18%	2183	0,20%
	10000	1204	0,20%	2092	0,20%
	20000	1287,2	0,16%	2135	0,24%
	30000	1298,9	0,25%	2235	0,27%
	40000	1325,4	0,13%	2089	0,30%
	50000	1240,1	0,17%	2185	0,35%
	60000	1189,4	0,16%	2210	0,44%
	70000	1144,2	0,16%	2188	0,23%
	80000	1270	0,14%	2335	0,054%
200	0	1040	0,32%	2789	0,40%
	10000	1099	0,34%	2478	0,33%
	20000	1178	0,14%	2984	0,41%
	30000	1278	0,59%	3174	0,57%
	40000	1394	0,47%	3299	0,59%
	50000	1662	0,24%	3487	0,40%
	60000	1802	0,78%	3399	0,30%
	70000	1877	0,98%	3719	0,74%
	80000	1947	0,23%	3879	0,650%

		Formulación Extendida		Progressive Hedging	
Escenarios	Demanda	Tiempo[s]	Gap [%]	Tiempo[s]	Gap [%]
300	0	2281	0,74%	4209	0,74%
	10000	2414	0,66%	4971	0,98%
	20000	2593	0,85%	5378	1,20%
	30000	2662	0,10%	5674	0,11%
	40000	2883	0,20%	5982	0,16%
	50000	3377	0,35%	6209	0,74%
	60000	3587	0,64%	6662	0,45%
	70000	3714	0,10%	6887	0,66%
	80000	3755	0,54%	6987	0,800%
400	0	4566	0,45	4789	0,78%
	10000	4891	0,75%	4991	0,99%
	20000	5401	1,50%	5364	1,42%
	30000	5589	1,60%	5472	0,45%
	40000	6984	1,54%	6477	0,67%
	50000	8412	1,59%	7888	0,49%
	60000	8556	1,67%	7936	0,98%
	70000	12478	1,03%	8985	0,74%
	80000	14289,1	0,87%	9005,3	0,80%
440	0	8693	0,69	8754	0,78
	10000	8745	0,78	8803	0,88%
	20000	16603	0,89	9421	0,65%
	30000	16993	0,77	9858	0,49%
	40000	17485	0,99	9745	0,89%
	50000	19114	1,24	9865	0,74%
	60000	19245	1,65	9888	1,30%
	70000	19878	1,25	9748	0,89%
	80000	20301,74	1,19%	10388	1,20%
500	0	11874	1,04	9974	1,24
	10000	12746	1,09	10004	1,22
	20000	15789	1,19	10115	1,04
	30000	18774	1,17	10875	0,98
	40000	20498	0,98	11365	0,97
	50000	24169	0,74	11497	1,57
	60000	26789	1,54	11999	1,69
	70000	27998	2	12785	1,87
	80000	28996,1	1,40%	13478	1,44%

		Formulación Extendida		Progressive Hedging	
Escenarios	Demanda	Tiempo[s]	Gap [%]	Tiempo[s]	Gap [%]
600	0	15784	1,45	12429	1,55
	10000	15996	1,89	12786	1,23
	20000	17489	1,77	12988	1,36
	30000	20165	2	12759	1,34
	40000	25412	1,98	13098	1,38
	50000	29555	1,56	13178	1,47
	60000	33761	1,35	13978	1,99
	70000	35474	1,2	14854	1,56
	80000	36875,1	1,79%	15669,9	1,78%
700	0	NO SOL	NO SOL	15894	1,88
	10000	NO SOL	NO SOL	16022	1,87
	20000	NO SOL	NO SOL	16424	1,86
	30000	NO SOL	NO SOL	16985	1,74
	40000	NO SOL	NO SOL	17469	1,77
	50000	NO SOL	NO SOL	17996	1,76
	60000	NO SOL	NO SOL	18743	1,94
	70000	NO SOL	NO SOL	20487	1,95
	80000	NO SOL	NO SOL	22157,4	1,94%
800	0	NO SOL	NO SOL	20141	1,77
	10000	NO SOL	NO SOL	21741	1,11
	20000	NO SOL	NO SOL	21984	1,24
	30000	NO SOL	NO SOL	21657	1,74
	40000	NO SOL	NO SOL	22436	1,65
	50000	NO SOL	NO SOL	23006	1,78
	60000	NO SOL	NO SOL	24789	1,45
	70000	NO SOL	NO SOL	27548	1,98
	80000	NO SOL	NO SOL	29545,2	1,89%
900	0	NO SOL	NO SOL	24885	1,27
	10000	NO SOL	NO SOL	25777	1,45
	20000	NO SOL	NO SOL	25987	1,66
	30000	NO SOL	NO SOL	25748	1,68
	40000	NO SOL	NO SOL	26098	1,56
	50000	NO SOL	NO SOL	29784	1,23
	60000	NO SOL	NO SOL	32479	1,47
	70000	NO SOL	NO SOL	33696	1,78
	80000	NO SOL	NO SOL	35488,2	2,01%

		Formulación Extendida		Progressive Hedging	
Escenarios	Demanda	Tiempo[s]	Gap [%]	Tiempo[s]	Gap [%]
1000	0	NO SOL	NO SOL	27464	2
	10000	NO SOL	NO SOL	28748	2
	20000	NO SOL	NO SOL	28996	2
	30000	NO SOL	NO SOL	28997	1,76
	40000	NO SOL	NO SOL	30451	1,64
	50000	NO SOL	NO SOL	34587	1,52
	60000	NO SOL	NO SOL	35664	1,89
	70000	NO SOL	NO SOL	39875	1,69
	80000	NO SOL	NO SOL	38541,2	2,54%
1240	0	NO SOL	NO SOL	30741	2
	10000	NO SOL	NO SOL	35987	2
	20000	NO SOL	NO SOL	36581	1,89
	30000	NO SOL	NO SOL	36665	1,88
	40000	NO SOL	NO SOL	37884	1,78
	50000	NO SOL	NO SOL	38456	1,74
	60000	NO SOL	NO SOL	39991	2
	70000	NO SOL	NO SOL	44251	1,98
	80000	NO SOL	NO SOL	43200	3,12%