

R E V I S T A

INGENIERIA DE SISTEMAS

Volumen XXII

Septiembre 2008

Sistema de Gestión de Personal para el Canal de Ventas de Metro S.A.	5
<i>Jaime Miranda P., Pablo A. Rey, Antoine Sauré V., Richard Weber, Cristóbal Montecino, José Mosquera</i>	
Mejoramiento de la Gestión y Uso de Tics en las Mipymes a Través de Software Libre	31
<i>Ángel Jiménez Molina, Juan D. Velásquez Silva, Alejandro Fuentes De La Hoz</i>	
Métodos cuantitativos para una nueva configuración de Unidades Territoriales en licitaciones de JUNAEB	57
<i>Gonzalo Zamorano, Guillermo Durán, Andrés Weintraub</i>	
Metodologías basadas en VaR para el análisis de riesgo en proyectos de embalses	85
<i>Prosper Lamothe, Eduardo Contreras</i>	
Sistema de Monitoreo de Mercado: Aplicación al Sector Eléctrico	109
<i>Rodrigo Palma Behnke, Sharo Escobar, Tomás Reid</i>	
Programación de Horarios de Clases y Asignación de Salas para la Facultad de Ingeniería de la Universidad Diego Portales Mediante un Enfoque de Programación Entera	123
<i>Rodrigo Hernández, Jaime Miranda P., Pablo A. Rey</i>	

Publicada por el

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

R E V I S T A
INGENIERIA DE SISTEMAS

ISSN 0716 - 1174

EDITOR

Guillermo Durán

Departamento de Ingeniería Industrial

Universidad de Chile

AYUDANTE DE EDICIÓN

Francisco Cisternas

Departamento de Ingeniería Industrial

Universidad de Chile

COMITÉ EDITORIAL

Sergio Maturana

Universidad Católica de Chile

Lorena Pradenas

Universidad de Concepción

René Caldentey

Universidad de Chile

Juan Carlos Sáez

JCSáez Editor

Rafael Epstein

Universidad de Chile

Ernesto Santibáñez

Universidad Católica de Valparaíso

Luis Llanos

CMPC Celulosa

Víctor Parada

Universidad de Santiago de Chile

Miguel Nussbaum

Universidad Católica de Chile

Jorge Vera

Universidad Católica de Chile

Financiado parcialmente por el Instituto Milenio “Sistemas Complejos de Ingeniería”, como reconocimiento a la difusión de las materias abordadas y de sus participantes.

Las opiniones y afirmaciones expuestas representan los puntos de vista de sus autores y no necesariamente coinciden con los del Departamento de Ingeniería Industrial de la Universidad de Chile.

Instrucciones a los autores:

Los autores deben enviar 2 copias del manuscrito que desean someter a referato a: Comité Editorial, Revista Ingeniería de Sistemas, Av. República 701, Santiago, Chile. Los manuscritos deben estar impresos en hojas tamaño carta, a doble espacio, deben incluir un resumen de no más de 150 palabras, y su extensión no debe exceder las 25 páginas. Detalles en www.dii.uchile.cl/~ris

Los artículos sólo pueden ser reproducidos previa autorización del Editor y de los autores.

Correo electrónico: ris@dii.uchile.cl

Web URL: www.dii.uchile.cl/~ris

Representante legal: Máximo Bosch

Dirección: República 701, Santiago, Chile.

Diagramación: Francisco Cisternas

Impresión: Ka2 Diseño e Impresión

Mail: contacto@ka2.cl

Carta Editorial Volumen XXII

Nos es muy grato editar este nuevo número de la Revista de Ingeniería de Sistemas (RIS) dedicado a temas de frontera en Investigación de Operaciones, Gestión y Tecnología. Queremos agradecer al Instituto Milenio “Sistemas Complejos de Ingeniería” por su colaboración para hacer posible esta publicación.

Este número contiene artículos de académicos y estudiantes de nuestro Departamento de Ingeniería Industrial (algunos de ellos incluso son consecuencia de trabajos finales de grado, tesis de magíster o tesis de doctorado), y de investigadores del Instituto Milenio.

La revista muestra trabajos sobre la aplicación de técnicas modernas de gestión de operaciones, programación matemática, tecnologías de información, finanzas y mercado eléctrico.

Nuestro objetivo a través de esta publicación es contribuir a la generación y difusión de las tecnologías modernas de gestión y administración. La revista pretende destacar la importancia de generar conocimiento en tecnología y administración para nuestras problemáticas, junto con adaptar las tecnologías foráneas a las realidades de nuestro país y de otros similares.

Estamos seguros de que los artículos publicados en esta oportunidad muestran formas de trabajo innovadoras que serán de gran utilidad e inspiración para todos los lectores, ya sean académicos o profesionales, por lo que esperamos que esta iniciativa tenga la recepción que creemos se merece.

Guillermo Durán
Editor

SISTEMA DE GESTIÓN DE PERSONAL PARA EL CANAL DE VENTAS DE METRO S.A.

JAIME MIRANDA P.*

PABLO A. REY**

ANTOINE SAURÉ V.***

RICHARD WEBER****

CRISTÓBAL MONTECINO*****

JOSÉ MOSQUERA*****

Resumen

Motivados por la creciente preocupación de la Gerencia Comercial de Metro S.A. respecto a los costos asociados a la gestión de su Canal de Ventas y por los nuevos estándares de calidad exigidos con la entrada en operación del nuevo plan de transporte de Santiago (Transantiago), se desarrolló un sistema computacional de gestión de personal que permite determinar, para cada instante del día y para cada uno de los puntos de venta de la red de Metro, la dotación de personal necesaria para satisfacer la demanda por venta de boletos. Lo anterior, mientras se tratan de satisfacer los estándares de calidad de servicio a mínimo costo. La metodología descrita en este trabajo reúne dos de los métodos más populares en la investigación de operaciones, como lo son programación entera y simulación, y se basa en la interacción de estas técnicas para poder incorporar las complejidades del sistema en estudio y conseguir mejores resultados. Su aplicación al problema de la Gerencia Comercial de Metro S.A. ha probado ser muy útil en cuanto al apoyo a las decisiones de programación de personal y ha traído numerosos beneficios en términos de reducción de costos, mejoramiento del nivel de servicio y automatización de procesos.

* Departamento de Control de Gestión y Sistemas de Información, Facultad de Economía y Negocios, Universidad de Chile, Santiago, Chile.

** Escuela de Ingeniería Industrial, Facultad de Ingeniería, Universidad Diego Portales, Santiago, Chile.

*** Operations and Logistics Division, Sauder School of Business, University of British Columbia, Vancouver, Canada.

**** Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.

***** Gerencia Comercial, Metro S.A., Santiago, Chile.

Palabras Clave: Programación de personal; sistema de soporte a las decisiones; programación matemática; simulación.

1. Introducción

En una ciudad con más de 5 millones de personas, el transporte público juega un rol importante para la calidad de vida de los habitantes. En ese sentido desde el año 1975, la empresa Metro S.A. ha ofrecido un servicio fundamental para el transporte de los residentes de Santiago de Chile. Sin embargo, desde la implementación del nuevo sistema de transporte Transantiago, en febrero del 2007, Metro S.A. se ha visto enfrentado a una serie de nuevos y complejos desafíos. Uno de ellos es la programación del personal de venta que debe satisfacer tanto la demanda por venta de cupos de transporte como las exigencias impuestas por el *Administrador Financiero del Transantiago* (AFT), todo esto a mínimo costo. El AFT, además de ser la institución encargada de recaudar y administrar los ingresos del Transantiago, es el encargado de custodiar y fiscalizar las operaciones de todos los agentes del sistema. En el caso de Metro S.A., el AFT fiscaliza el máximo número de usuarios en fila esperando por la venta de cupos de transporte, y el tiempo máximo de espera por este servicio.

Metro S.A. cuenta con una red de transporte compuesta por 5 líneas, 90 estaciones y 102 mesaninas¹, en sus 3 secciones². En esta red se desplazan más de 2.300.000 usuarios diariamente (noviembre 2007).

Entre las decisiones operacionales que tiene que tomar la *Gerencia Comercial* de Metro S.A. se encuentra la programación del servicio de caja para las distintas mesaninas de cada estación. En el contexto del presente artículo, se entiende como operario de caja a un vendedor de boletos asignado a un puesto fijo dentro de una boletería o a un vendedor fuera de una boletería (también llamado “rompe fila”).

De esta manera, la programación de personal se traduce en la definición, en cuanto a número y composición, de los requerimientos de personal de venta para cada momento de operación del servicio. Estos requerimientos son posteriormente entregados a los distintos operadores para su implementación.

¹**Mesanina:** Es el lugar físico donde se encuentran las boleterías y los torniquetes de acceso a los andenes en una estación de Metro. Cada mesanina puede tener dos tipos de puestos de venta en sus boleterías: con POS y sin POS. Los puestos de venta sin POS sólo pueden realizar ventas de boletos, en cambio los puestos de venta con POS pueden atender tanto usuarios que desean comprar boletos como aquellos que desean cargar su tarjeta multivía. Los usuarios que desean cargar su tarjeta multivía solamente pueden hacerlo en un puesto de venta con POS.

²**Sección:** Corresponde a un conjunto de mesaninas asociadas a un operador. Un operador es un empresa que provee servicio de personal a Metro S.A.

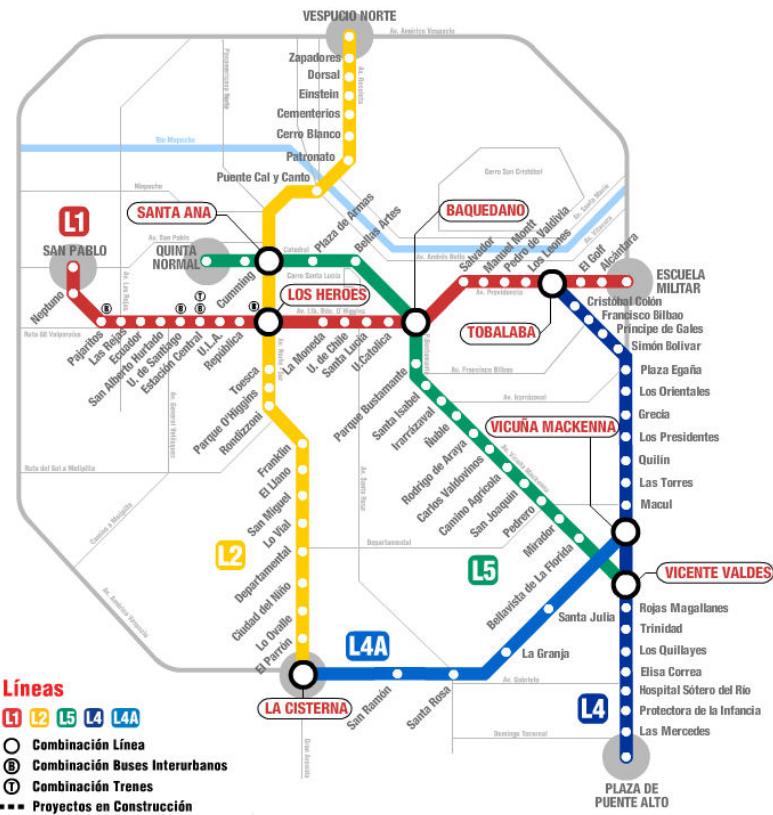


Figura 1: Red de líneas y estaciones de Metro S.A.

En este artículo se presenta una metodología para abordar el problema de programación, enfrentado por la *Gerencia Comercial* de Metro S.A., que mezcla optimización y simulación. Esta metodología ha sido implementada en un sistema computacional que genera la programación de servicios de caja en forma automática y busca el cumplimiento de los niveles de servicio definidos por la *AFT*, mientras se minimizan los costos. El sistema está compuesto por un *Módulo de Asignación óptima* y un *Módulo de Simulación* que interactúan iterativamente.

Actualmente, la programación de los cajeros se realiza a través del juicio de los expertos del negocio mediante prueba y error. La programación de cada mesanina es realizada sobre la base de un solo día “*tipo*”, replicando esta misma programación de operarios para todos los días del mes. El principal problema de este procedimiento es que no considera las fluctuaciones de la demanda durante el transcurso de un día, ni tampoco a través de los días de una semana.

Como se mencionó anteriormente, la programación final de operarios es entregada a un conjunto de empresas privadas, llamadas *operadores*, que se encargan de satisfacer los requerimientos definidos por Metro S.A. Como con-

secuencia, para la *Gerencia Comercial* de Metro S.A. los operarios/cajeros son indistinguibles, sólo diferenciándolos por su tipo y el largo de sus turnos.

Otro antecedente relevante para este problema es que el número de estaciones de la red ha ido en aumento de manera sostenida (ver Figura 2), haciendo de la programación de operarios un problema aún más complejo. Adicionalmente, si consideramos que las decisiones de programación de personal se deben tomar de manera rápida y eficiente, se realza aún más la necesidad de constar con un sistema que apoye las decisiones estudiadas en este trabajo.



Figura 2: Evolución del número de estaciones de Metro S.A. en el tiempo.

El sistema computacional propuesto permite a Metro S.A. obtener una programación óptima de los operarios de caja ajustándose a los diferentes comportamientos de la demanda para cada mesanina de la red.

El presente artículo tiene la siguiente estructura: la sección 2 presenta los antecedentes relevantes para este trabajo. Primero se expone el estado-del-arte en relación al problema analizado. Luego se realiza una descripción sistemática del proceso de gestión de personal, caracterizando cada uno de sus subprocesos, reglas de operación y recursos. La sección 3 detalla los modelos que conforman la base del sistema desarrollado. La sección 4 analiza los beneficios de la metodología propuesta, en términos de costos, violaciones de los estándares de servicio y automatización del proceso de toma de decisiones. Finalmente, la sección 5 muestra las principales conclusiones del proyecto.

2. Antecedentes

2.1. Estado del Arte

El problema de programación de operarios requeridos en un sistema ha sido ampliamente estudiado en la literatura a medida que las empresas se han orientado a mejorar su nivel de servicio y más conscientes de sus costos. Este interés ha abarcado desde sistemas muy simples hasta sistemas muy complejos. En la teoría, este problema puede verse como un subgrupo de un conjunto más general de problemas llamado *Personnel Scheduling* que considera tanto la programación como la asignación de personas, tareas, turnos, rutas y tripulaciones, entre otros [13, 14].

La programación de personal considera la determinación de los niveles óptimos de personal para satisfacer una demanda que puede ser por uno o más servicios y que puede variar significativamente con el transcurso del tiempo. Lo anterior, mientras se satisfacen ciertos estándares de calidad. Este problema fue inicialmente discutido en los años 50 por Edie [12] y Dantzig [10]. En 1954, Edie [12] propuso el uso de probabilidades para calcular el número de cajetas y, por lo tanto, del número de cajeros, requeridas para proveer un nivel de servicio determinado a diferentes horas del día en una plaza de peaje. Dantzig [10] por su parte esbozó el uso de programación lineal para programar dichos operadores.

Desde entonces, los problemas de programación de personal han sido de gran interés tanto en el ámbito comercial como en el ámbito académico. Esto ha permitido el desarrollo de numerosas aplicaciones en áreas que, principalmente en la industria de servicios, se han caracterizado por requerir el manejo de un gran número de personas y una demanda que varía enormemente durante el día. Algunas de estas áreas son *call centers* [3, 8], correos [22], universidades [2], servicios de emergencia [7], aeropuertos [20], servicios de salud [1, 7, 11], entre muchas otras.

En el ámbito académico, la literatura de programación de personal puede ser clasificada en dos grandes grupos. El primer grupo reúne los trabajos que utilizan métodos tradicionales de solución como programación lineal o entera y simulación. El segundo grupo, en cambio, reúne todos los trabajos que utilizan enfoques especializados de solución y que son creados para problemas particulares. Nuestro trabajo se encuentra en el primer grupo.

Lo novedoso de nuestro estudio es que si bien hace uso de dos de los métodos más populares en la investigación de operaciones – como lo son programación entera y simulación – se basa en la interacción de estas técnicas

para poder incorporar las complejidades del sistema en estudio y conseguir mejores resultados.

En los últimos 10 años, gracias a la mayor rapidez y menores costos proporcionados por el desarrollo de los computadores y los programas de simulación, este enfoque de optimización y simulación se ha hecho muy popular, siendo ampliamente utilizado en diferentes sectores. Por ejemplo, se ha empleado para resolver diversos tipos de problemas en bancos [15], correos [22], aeropuertos [20], hospitales [7], servicios de *courier* [9] y sistemas productivos [6, 19, 17], entre otros [5, 21]. Su aplicación a problemas de programación de personal en *call centers* se ha hecho de particular interés dada la fluctuante naturaleza de la carga de trabajo en estos sistemas y de lo exigente de sus niveles de servicio (dos factores no tan diferentes a los estudiados en este proyecto) [3, 8, 16].

Azadivar [4] describe el problema de optimización-simulación o simulación-optimización básicamente como un problema de optimización donde la función objetivo y algunas restricciones sólo pueden ser evaluadas usando simulación. En este contexto, la simulación es utilizada para probar la validez del modelo o para generar las entradas para una nueva instancia del problema de optimización. Esta interacción permite disminuir drásticamente el número de escenarios a ser analizados mientras se alcanza una mejor solución en términos de costos y niveles de servicio [5, 18].

Para terminar esta sección es importante mencionar que los trabajos publicados por Ernst *et al.* [13, 14] proveen una rigurosa revisión bibliográfica de todos los tipos de problemas relacionados con programación de personal, clasificados por área de aplicación y algoritmos de solución. Además, estos trabajos incluyen secciones dedicadas exclusivamente a artículos sobre modelamiento de la demanda y pronósticos, actividades necesarias para definir adecuadamente los requerimientos de personal.

2.2. Descripción de Problema

La *Gerencia Comercial* de Metro S.A., a través de su Canal de Ventas, tiene la misión de proveer a los usuarios de la red de Metro S.A. con cupos de transporte. Este objetivo se lleva a cabo mediante la programación de los operarios de ventas en las distintas mesaninas de la red.

Cada 15 días, la *Gerencia Comercial* de Metro S.A. define los requerimientos de personal para toda red, estableciendo el número de operarios y turnos para cada uno de sus tipos de cajero. La *Gerencia Comercial* incorpora dos criterios. El primer criterio corresponde a la *rentabilización* del servicio mediante la minimización de los costos operacionales. El segundo corresponde al cumplimiento de los niveles mínimos de servicio definidos por el *AFT*.

Los estándares mínimos definidos por el *AFT* están relacionados con el largo máximo de usuarios en fila y el tiempo de espera máximo durante una

hora móvil.

2.3. Proceso Global de Planificación y Control de los Operarios de Caja

El proceso global de planificación y control del servicio de venta de cupos de transporte está compuesto por 4 etapas. La Figura 3 muestra dicho proceso.

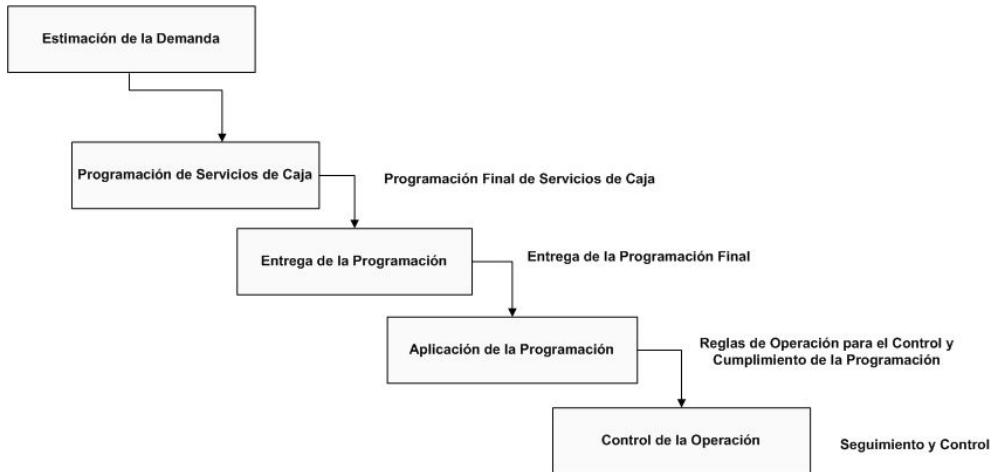


Figura 3: Etapas del proceso global de planificación y control de operarios.

Este proceso comienza con la *Estimación de la Demanda* por el servicio de venta, la que se define como el número de usuarios que llegarán a una mesanina solicitando la compra de un boleto o carga de su tarjeta multivía para cada intervalo de 15 minutos³. Luego esta estimación sirve como entrada para la *Programación de Operarios de Caja*. La programación final define el número, la hora de inicio y la hora de término de cada turno para cada tipo de operario. La siguiente etapa corresponde a la *Entrega de la Programación Final a los Operadores*. Como se mencionó anteriormente, estos *operadores* son empresas privadas responsables de proveer los operarios de ventas de acuerdo a los requerimientos definidos por Metro S.A. Posteriormente, se hace la *Aplicación de la Programación* en la red. Finalmente, se realiza el *Control de la Operación*, donde se vela por el cumplimiento de los requerimientos de personal y turnos definidos por Metro S.A.

³La estimación de demanda es entregada por la *Gerencia Comercial* obtenida en un proceso externo al sistema.

3. Diseño del Sistema Computacional

Para apoyar las decisiones de programación de operarios de caja se diseñó y construyó un sistema computacional. Este sistema computacional está compuesto por dos módulos integrados. El primer módulo, al que llamamos *Módulo de Asignación*, genera una programación óptima de operarios de caja para todas las mesaninas de la red, suponiendo todos los parámetros en sus valores promedios. Entre estos parámetros tenemos los tiempos de atención y la demanda por boletos y cargas de tarjeta multivía.

El segundo módulo, al que llamamos *Módulo de Simulación*, tiene como objetivo la evaluación y corrección de la programación definida por el *Módulo de Asignación*. Esta evaluación se realiza simulando el funcionamiento del *Canal de Ventas* bajo la programación de operarios definida por el *Módulo de Asignación*, esta vez incorporando una descripción probabilística más detallada de los procesos de llegada de usuarios y de atención. Lo anterior otorga mayor realismo a los resultados del sistema computacional.

Un elemento interesante de destacar de la metodología propuesta en este trabajo es la interacción existente entre los módulos para obtener la programación final de operarios. Esta interacción comienza cuando el *Módulo de Asignación* genera una programación de operarios para la red. Luego dicha programación es simulada obteniendo para cada mesanina distintos indicadores de desempeño y niveles de servicio. Si estos indicadores no cumplen con los estándares mínimos exigidos por el *AFT*, el *Módulo de Simulación* genera una serie de nuevas restricciones, las cuales son incorporadas al *Módulo de Asignación*. Con estas nuevas restricciones, se vuelve a generar una programación y ésta se valida nuevamente mediante su simulación. Este proceso se repite hasta que se cumplan todos los estándares mínimos de operación. La Figura 4 ilustra el proceso iterativo de ajuste para la generación de la programación final de servicios.

A continuación, se hace una descripción detallada de las funcionalidades y características de cada módulo.

3.1. Módulo de Asignación

Este módulo apoya las decisiones de determinación del número, hora de inicio y hora de término para cada tipo de cajero considerado por Metro S.A.. Su propósito es la minimización de los costos totales asociados a la programación de personal, mientras se satisfacen los requerimientos definidos por el *AFT* y las condiciones deseables de operación del sistema.

Este módulo recibe como información de entrada las estimaciones de de-

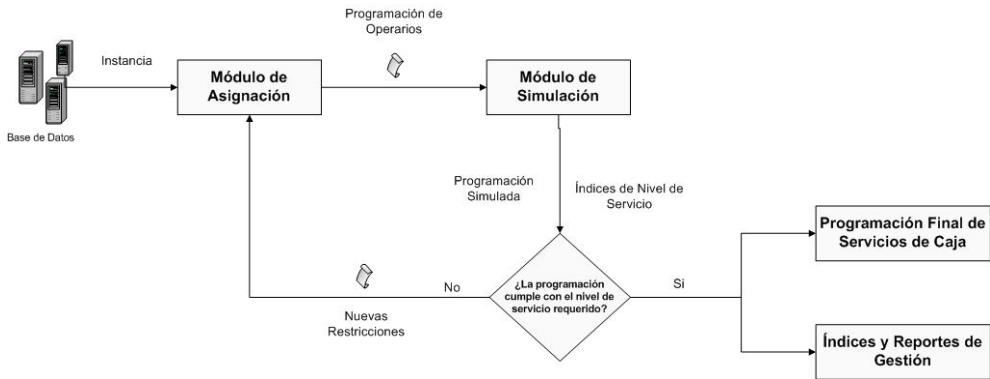


Figura 4: Etapas del proceso de programación de operarios de caja.

manda elaboradas actualmente por la *Gerencia Comercial*; la configuración de la red de Metro S.A. en términos del número de líneas, estaciones y mesaninas; la cantidad de recursos disponibles por cada mesanina (como por ejemplo el número de *POS*); y las exigencias impuestas en los contratos con el *AFT*.

3.1.1. Descripción del Módulo de Asignación

A continuación, se presenta una descripción más detallada del *Módulo de Asignación* del sistema. Este módulo se basa en un modelo de programación entera e implementado en el software GAMS 22.5 y resuelto con el paquete CPLEX 10.0.

Función Objetivo

El objetivo de nuestro modelo es minimizar el costo total asociado a la programación de operarios. Dado que no existe ninguna relación que ligue una mesanina con otra en términos operacionales, nuestro problema se reduce a minimizar el costos total para cada una de las mesaninas por separado.

Parámetros del Módulo

1. *Demandas de Usuarios*: Corresponde al número de usuarios que llegarán solicitando el servicio de caja en cada intervalo de 15 minutos en el horizonte de programación. Esta demanda se estima para cada mesanina de la red.
2. *Configuración de las Mesaninas*: Corresponde a la cantidad de ventanillas operativas y *POS* disponibles para el servicio de venta.
3. *Tiempo de Atención*: Corresponde al tiempo que tarda un operario en

atender una solicitud de un usuario. Este tiempo fue considerado para cada tipo de transacción.

4. *Costos:* Los costos asociados a la programación de operarios. Estos costos difieren dependiendo tanto del número como del largo de los turnos.

Requerimientos y Condiciones del Problema

Los requerimientos que la programación final de operarios debe satisfacer. Estos se relacionan principalmente a las diferentes exigencias impuestas para la generación de la programación de operarios. La formulación matemática de estos requerimientos se puede encontrar en el Anexo.

Las condiciones y requerimientos del sistema son los siguientes:

1. *Número Máximo de Operarios:* Esta restricción impone que se respete el número máximo de puestos de venta que tiene cada mesanina.
2. *Criterio de Habilitación de Puestos Móviles:* Esta restricción impone que antes de habilitar un puesto móvil, es necesario habilitar primero todos los puestos fijos de una mesanina.
3. *Cantidad Mínima de Operarios:* Esta condición impone que para cada mesanina-período debe existir una cantidad mínima de operarios.
4. *Relación entre Tasas de Llegada y Atención:* Esta restricción impone que la tasa total de atención en una mesanina debe ser mayor o igual que la tasa de llegada de usuarios a esta. La tasa de atención depende directamente de la cantidad de operarios activos y la tasa de llegada depende de la estimación de demanda en intervalo de tiempo.
5. *Eliminación de Ventanas Temporales:* Esta restricción trata de que el número de operarios requeridos no varíe bruscamente entre períodos de tiempo consecutivos. .
6. *Restricción de Corrección del Módulo de Simulación:* Consiste en una o más restricciones, provenientes del *Módulo de Simulación*, que se imponen para cada intervalo de tiempo que no cumple con el nivel mínimo de servicio requerido. Esta restricción establece que la cantidad total de operarios sea mayor en una unidad respecto de la cantidad total de servidores que se programó en la iteración anterior del sistema.

3.2. Módulo de Simulación

A continuación, se presenta una descripción más detallada del *Modulo de Simulación* del sistema. Este módulo fue implementado en Java System Development version 6.3 para Sun Microsystems.

3.2.1. Descripción del Proceso a Simular

Los usuarios llegan a una mesanina a comprar boletos, a cargar una tarjeta multivía o por otras actividades no transaccionales. Cuando un usuario está frente a la mesanina debe escoger una fila para ubicarse y, en el caso en que no haya fila, escoger un operario (ventanilla) para ser atendido inmediatamente. Si un usuario necesita cargar su tarjeta multivía debe seleccionar un operario que esté habilitado para brindar este servicio, es decir, debe seleccionar un operario que cuente con *POS*.

Un usuario que no es atendido inmediatamente avanza en su posición relativa dentro de la fila, a medida que los usuarios que están antes que él van siendo atendidos, hasta llegar al operario de venta. En ese momento se realiza la transacción, la cual tiene una duración variable que depende del tipo de transacción. Una vez realizada la transacción el usuario sale del sistema.

La Figura 5 muestra un diagrama que identifica la entrada, salida, reglamentos y recursos que dispone el sistema:

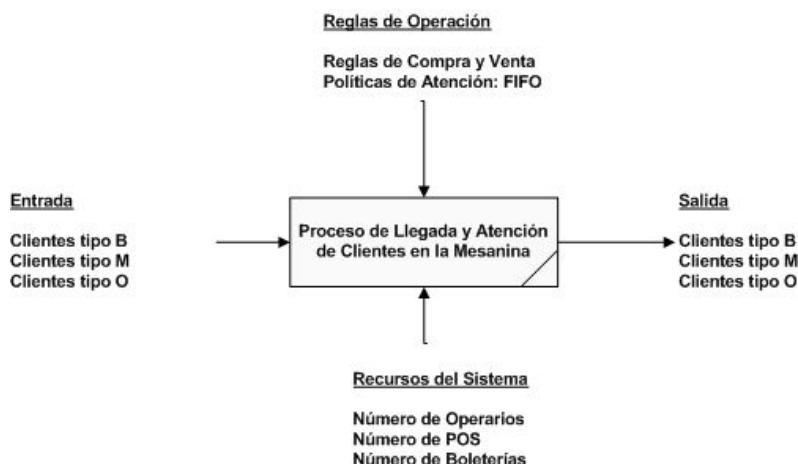


Figura 5: Diagrama IDEF0 para el *Módulo de Simulación*.

Los elementos de la figura anterior se describen a continuación:

- **Entrada:** La entrada del sistema corresponde a los usuarios de la red. Estos usuarios pueden ser de tres tipos, usuarios que compran boletos, usuarios que cargan tarjetas multivía y usuarios que realizan otras actividades no transaccionales. Un ejemplo de estos últimos corresponde a los clientes que hacen consultas por tramos especiales, consultas sobre las estaciones que son conexiones de línea o terminales de buses.
- **Salida:** La salida del sistema corresponde a los usuarios que ya compraron boletos o cargaron su tarjeta multivía.

- *Reglas*: Las reglas impuestas corresponden a restricciones operacionales impuestas al sistema. Por ejemplo, la política de atención de los operarios (FIFO) y las políticas de compra y venta de boletos (los usuarios que deseen cargar su tarjeta multivía se deben ubicar en filas que tengan POS).
- *Recursos*: Los principales recursos utilizados corresponden a los operarios de caja, tanto fijos como móviles. Además está el espacio físico dónde se realizan las transacciones (i.e. boleterías). A esto se suman todos los equipos necesarios, como son por ejemplo: número de *POS*, boletos, entre otros.

3.2.2. Descripción Lógica del Proceso

En términos generales, la simulación queda bien representada por los eventos que se muestran en la Figura 6.



Figura 6: Esquema general del *Módulo de Simulación*.

A continuación, se detalla cada una de estas etapas:

1. Etapa 1 - Proceso de Llegada

Se considera que las llegadas de los usuarios al sistema son individuales e independientes, es decir, los usuarios no llegan en grupo, y la probabilidad de que dos o más usuarios ingresen al mismo tiempo al sistema es cero.

En el proceso de llegada se distinguen dos variables aleatorias independientes. La primera de ellas corresponde a los *Tiempos Entre Llegadas* sucesivas de los usuarios. El *Módulo de Simulación* considera distintas distribuciones de probabilidades para definir estos tiempos.

- Caso Determinístico*: Considera que los tiempos entre llegadas de usuarios sucesivos corresponden a intervalos fijos de tiempo.
- Caso Probabilístico*: Considera que los tiempos entre llegadas son posibles de ajustar mediante algún tipo de distribución de probabilidades. Las distribuciones de probabilidades utilizadas fueron: Exponencial, Normal, Lognormal, Weibull y Uniforme.

En segundo lugar están los tipos de usuario. La forma de abordar esta situación fue mediante la incorporación de una *Probabilidad* asociada a *Tipo de Transacción*.

2. Etapa 2 - Elección del Operario (Fila)

Cuando un usuario llega a una mesanina debe decidir en que fila ponerse. La forma en que los usuarios eligen la fila (u operario) tiene incidencia directa en los indicadores de desempeño del sistema y, por lo tanto, afectan directamente a la programación final del personal. Debido a esto, el *Módulo de Simulación* debe permitir utilizar distintas reglas de elección de fila. Cabe destacar que para los efectos de la simulación se considera que un usuario ya en una fila posteriormente no se cambia, y en el caso en que dos operarios tengan el mismo largo de fila, los usuarios eligen equiprobablemente. Los tipos de elección implementados son los siguientes:

- a) *Elección Equiprobable*: Los usuarios que llegan al sistema tienen igual probabilidad de elección entre los distintos operarios. Los usuarios que requieren cargar una tarjeta multivía sólo lo pueden hacer con un operario que cuente con POS. En cambio los usuarios que compran boleto lo pueden hacer con cualquier operario activo.
- b) *Elección Operario con Fila de Menor Tamaño*: Dentro de este tipo de elección podemos definir tres casos:
 - 1) *Asignación Directa al Operario con Fila más corta*: Los usuarios que compran boletos o que realizan actividades no transaccionales eligen directamente el operario con la fila más corta (pueden utilizar cualquier operario). Lo mismo ocurre con los usuarios que carga su tarjeta multivía. Sin embargo, estos últimos sólo consideran los operarios con *POS*.
 - 2) *Mayor probabilidad a la fila más corta*: Los usuarios que llegan al sistema eligen con mayor probabilidad al operario con la menor fila.
 - 3) *Mayor probabilidad al operario con fila más corta, pero los usuarios que compran boleto eligen con mayor probabilidad un servidor sin POS*: Este caso es similar al anterior, sin embargo, aquí se considera que existe un incentivo para que los usuarios que compran boleto seleccionen con mayor probabilidad una fila en donde el operario no cuenta con *POS*.

3. Etapa 3 - Atención de Caja

Se considera que las atenciones de los usuarios son individuales e independientes, es decir, no se atienden grupos. En el proceso de atención se distinguen dos eventos aleatorios independientes. El primero de ellos corresponde a los *Tiempos de Atención del Operario* el cual fue diferenciado para cada tipo de usuario del sistema. Dentro de este contexto, se consideraron dos casos:

- a) *Caso Determinístico*: Aquí se considera que los tiempos de atención de los usuarios son iguales.
- b) *Caso Probabilístico*: Este caso considera que los tiempos de atención se ajustan según una distribución de probabilidades. Las distribuciones de probabilidades utilizadas fueron las siguientes: Exponencial, Normal, Lognormal, Weibull y Uniforme.

4. Etapa 4 - Salida del Cliente

Esta etapa representa el término del servicio. Se considera que después de la venta o carga de tarjeta el usuario abandona el sistema, es decir, o hace uso del servicio de transporte o deja la mesanina.

4. Beneficios del uso del sistema computacional

Esta sección presenta una comparación de las programaciones obtenidas con la aplicación del sistema frente a las utilizadas por Metro S.A. De esta manera, se pretende evidenciar los beneficios que produce el uso del sistema computacional en términos cuantitativos. Los análisis presentados en esta sección corresponden a la comparación de los resultados de un *Plan Piloto* desarrollado en el mes de noviembre del año 2006.

Además, esta sección presenta una evaluación económica de los costos totales en que se incurre por cada programación para luego analizar el comportamiento de la solución entregada en términos del nivel de servicio provisto. También se incluye una comparación de la solución obtenida en términos del número de servicios de caja que se programan con respecto a la demanda. Este análisis permite observar cuán eficientes son nuestras programaciones.

Finalmente, se describen otros otros beneficios obtenidos con la aplicación del sistema desde el punto de vista de la organización.

4.1. Descripción del Plan Piloto

El objetivo central de la realización del *Plan Piloto* fue validar la utilidad de las programaciones generadas por el sistema computacional y analizar su desempeño respecto a las programaciones tradicionalmente utilizadas por la *Gerencia Comercial* de Metro S.A.

Para evaluar los beneficios de nuestro modelo se midió un número reducido de mesaninas (cercano al 10 % del total) durante cuatro días distribuidos en dos semanas. Las pruebas experimentales se realizaron durante la segunda y tercera semanas del mes de noviembre de 2006.

Se decidió utilizar para una de las semanas la programación propuesta por el sistema computacional, y para la programación confeccionada por Metro S.A. Las pruebas se realizaron los días martes y viernes, con la finalidad de minimizar la influencia de factores ajenos al sistema en estudio. Durante la primera semana se aplicó la programación del sistema computacional en la mitad de las mesaninas escogidas, y en el resto de las mesaninas, se aplicó la programación proporcionada por Metro S.A. En la segunda semana, se invirtió la selección de las programaciones.

Para realizar este *Plan Piloto* se escogieron 8 mesaninas, buscando definir una muestra representativa del total de las mesaninas que componen el sistema. El calendario del *Plan Piloto* se resume en la Tabla 1. La primera columna corresponde a la mesanina donde se realizó la medición. Las demás columnas indican si para dichas estaciones se utilizó la programación propuesta por el sistema o la programación propuesta por Metro S.A. Por ejemplo, para la estación Pajaritos (PJ) el día 14/11 se utilizó la programación de Metro S.A., y para la misma estación el día 21/11 se utilizó la programación del sistema.

Mesanina	14/11/2006	17/11/2006	21/11/2006	24/11/2006
Los Leones (LE)	Sistema	Sistema	Metro S.A.	Metro S.A.
Puente Alto (PPA)	Sistema	Sistema	Metro S.A.	Metro S.A.
Pedro de Valdivia (PV)	Sistema	Sistema	Metro S.A.	Metro S.A.
Universidad de Santiago (US)	Sistema	Sistema	Metro S.A.	Metro S.A.
Elisa Correa (ECO)	Metro S.A.	Metro S.A.	Sistema	Sistema
Estación Central (EL)	Metro S.A.	Metro S.A.	Sistema	Sistema
Escuela Militar Oriente (EM-Oriente)	Metro S.A.	Metro S.A.	Sistema	Sistema
Pajaritos (PJ)	Metro S.A.	Metro S.A.	Sistema	Sistema

Cuadro 1: Calendario de ejecución del Plan Piloto.

Para evaluar y comparar las programaciones se midieron los *largas de fila*, los *niveles de demanda* y los *tiempos de transacción*. Estas mediciones se realizaron en todas las mesaninas en estudio durante el período completo de operación, es decir, para cada día entre las 6:30 y las 22:30 horas. Para determinar los largos de fila se registró, minuto a minuto, cuántas personas estaban esperando a ser atendidas por cada ventanilla o rompefila designado. La demanda enfrentada fue determinada de la manera que usualmente lo realiza Metro S.A., a partir de información agregada extraída de sus sistemas. Específicamente, a partir de los registros, cada 15 minutos, del paso de usuarios a través del conjunto de torniquetes y del número total de transacciones de carga de tarjetas. Los tiempos de transacción fueron medidos directamente en un subconjunto de boleterías en cada mesanina.

A continuación, se describen los resultados del estudio y los beneficios generados por el sistema propuesto.

4.2. Beneficios cuantitativos

4.2.1. Comparación del ajuste de la programación a la demanda

En esta sección se compara la programación obtenida con el sistema propuesto y la programación definida por la *Gerencia Comercial* con su actual forma de operar. Esta comparación se realiza con respecto al número de operarios de caja requeridos y a la demanda estimada normalizada. Para esta comparación se utilizaron los datos correspondientes a la segunda semana del mes de noviembre (primera semana de implementación del Plan Piloto).

Las Figuras 7 y 8 muestran el número de operarios habilitados por cada programación, en cada bloque horario, y el nivel de demanda estimado dividido por 100. Este último valor es una estimación del mínimo de operarios de caja que se necesitan para satisfacer la demanda⁴. La intuición detrás de esta comparación es que, por un lado, una programación que consigue atender la demanda debe estar por encima de la curva de “Demanda/100”, y por otro, mientras menor sea el requerimiento de servicios de caja, más económica resulta la programación. Es decir, una buena programación debe estar por encima de la curva de demanda, pero debe seguirla lo más ajustadamente posible.

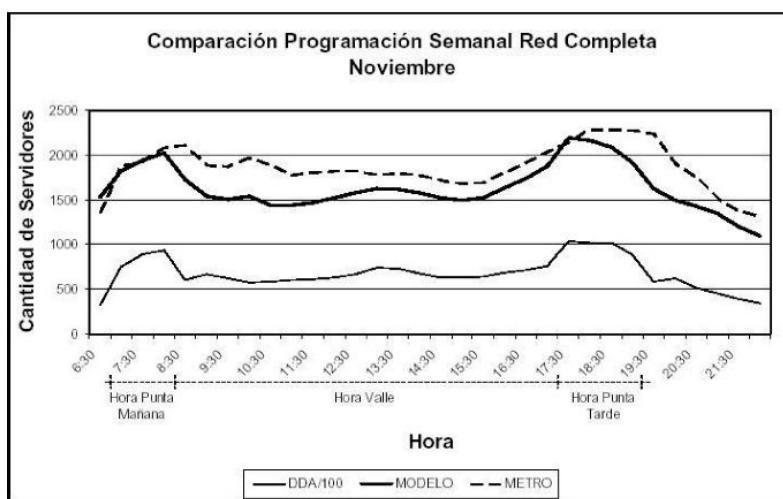


Figura 7: Comparación del número de servicios de caja asignados y la demanda real por servicios (red completa).

La Figura 7 muestra esta información, de manera agregada, para toda la red y considerando la programación de la semana completa. Es fácil ver que la planificación propuesta por el sistema se ajusta mejor a la curva de demanda normalizada. Es importante señalar que el área entre las curvas correspondientes a las dos programaciones representa, de manera aproximada, el ahorro

⁴Esto corresponde al caso en que los clientes llegan de manera homogénea y los tiempos de transacción son constantes.

en horas-hombre producido por el sistema computacional. Se observa que este ahorro proviene principalmente de la disminución de personal requerido durante las *horas valle*⁵ y luego de la *punta tarde*⁶.

Como ilustración de lo que sucede en una mesanina, la Figura 8 muestra el resultado para el día viernes en particular para la mesanina de la estación Los Leones. En este gráfico también se observa que la programación producida a partir de nuestros modelos se ajusta mejor a la demanda. En particular, la programación producida con el sistema computacional requiere de una menor cantidad de servicios de caja durante las horas valle. En cambio, en el horario punta más pronunciado se programan más cajeros que lo que hacía Metro S.A. con anterioridad al uso del sistema, pero por un período más corto de tiempo. De esta manera, con la programación generada por el sistema se atiende la demanda que crece momentáneamente pero manteniendo menores niveles de personal durante todo el día.

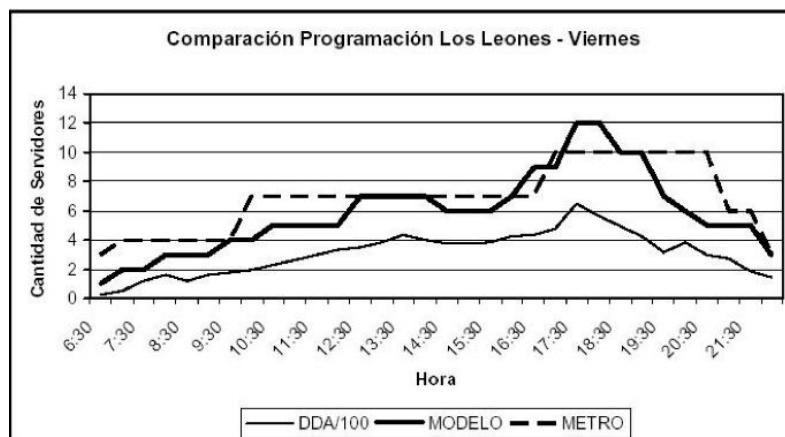


Figura 8: Comparación del número de servicios de caja asignados y la demanda real por servicios (mesanina Los Leones).

En conclusión, las programaciones generadas por el sistema computacional, en general, se aproximan más a la demanda real, aprovechando entonces mejor los recursos de operarios de caja.

4.2.2. Evaluación económica en términos del costo total y comparación respecto al indicador de nivel de servicio

En esta sección se comparan las dimensiones de costo y nivel de servicio para las programaciones generadas por el sistema computacional y por Metro S.A. Como indicadores del nivel de servicio se utilizaron: el número de

⁵Se refiere a las horas del día con baja demanda.

⁶Se refiere a las horas de baja demanda después de las 19:30 horas.

mesaninas-período⁷ en que el largo de fila promedio superó el estándar establecido de 3 usuarios; y el largo de fila promedio cuando se superó este estándar. Para realizar la comparación, se consideraron cuatro programaciones, a las cuales nos referimos como:

- Sistema 3 Días: Se generan tres programaciones distintas, una para los *días laborales*, otra para el sábado y, la tercera, para el domingo.
- Sistema 4 Días: Se generan cuatro programaciones distintas: una para lunes y martes; otra para miércoles, jueves y viernes; otra para el sábado; y la última para el domingo.
- Sistema 5 Días: Se generan cinco programaciones distintas: una para el lunes; otra para martes, miércoles y jueves; otra para el viernes; otra para el sábado; y la última, para el domingo.
- Sistema 7 Días: Se generan siete programaciones distintas, una para cada día de la semana.

La programación de servicios de caja se define para cada tipo de día, es decir, una misma programación para todos los días del mismo tipo. Para determinar la demanda en un determinado período del día tipo, se considera el máximo de la demanda estimada para todos los días de ese tipo en la semana. Por ejemplo, para determinar la demanda del período 8:30 a 8:45 horas de un día tipo “martes-miércoles-jueves” para la programación “Sistema 5 Días” se toma el máximo de las demandas estimadas para el martes, miércoles y jueves de 8:30 a 8:45 horas.

En todos los casos, para la generación de la programaciones, se utilizó como información de entrada las demandas estimadas por Metro S.A. para una semana tipo del mes de noviembre de 2006.

El comportamiento de estas cinco programaciones fue simulado utilizando la demanda real de la segunda semana de noviembre de 2006.

La Tabla 24.2.2 se muestra, para cada programación, el costo total (en UF), la variación porcentual de este costo respecto del caso base (programación Metro S.A.), el largo promedio de las filas en los casos que superan 3 personas y la proporción que este promedio representa respecto del caso base (programación Metro S.A.).

De la tabla anterior, se puede decir que, para el caso que considera tres tipos de día, el sistema genera una programación ligeramente más costosa que la utilizada por Metro S.A., pero reduciendo significativamente los largos de fila en los casos que se viola el estándar de servicio. Cabe señalar que para la programación actual Metro S.A. se ocupan 4 tipos de días

⁷Llamamos *período* a cada intervalo de 15 minutos utilizados como unidad mínima de tiempo para las mediciones y programaciones.

Programación	Costo (UF)	Variación Caso Base	Largo Promedio Filas Largas	Proporción Caso Base	Mesaninas-Período Fuera de Estándar	
					Número	Porcentaje
Real Metro	3.324,8	0,00 %	92,7	100,0 %	3.941	9,32 %
Sistema 3 Días	3.367,9	+1,29 %	13,5	14,6 %	535	1,27 %
Sistema 4 Días	3.233,1	-2,76 %	18,9	20,4 %	882	2,09 %
Sistema 5 Días	3.175,2	-4,50 %	21,2	22,9 %	1.145	2,71 %
Sistema 7 Días	3.075,0	-7,52 %	22,4	24,2 %	1.723	4,08 %

Cuadro 2: Comparación de programaciones para el mes de noviembre de 2006.

Para las otras programaciones se obtienen mejoras tanto en los costos como en la calidad de servicio. En particular, el largo promedio de las colas “largas” se reduce a menos de una cuarta parte de lo registrado actualmente en todos los casos.

4.3. Beneficios cualitativos

4.3.1. Automatización

La forma en que se programan los puestos de venta ha variado con los contratos que han regido para cada línea de la red de Metro S.A.. Sin embargo, este proceso ha mantenido características comunes, lo que permite describir una metodología estándar de programación.

Existe una programación base, que debe especificar los requerimientos mínimos con los que cada mesanina puede operar manteniendo un nivel de servicio aceptable. Cuando se detecta que esta base presenta falencias en algún punto de la red, se genera una solicitud para modificarla de manera eventual o permanente. Esta solicitud es evaluada y, en caso de determinarse pertinente, es puesta en marcha.

Esta manera de programar presenta deficiencias, que pueden ser en parte corregidas por el sistema computacional desarrollado, obteniéndose los siguientes beneficios:

- **Unificación del nivel de servicio.** El uso del sistema computacional permite obtener una programación que cumpla con un estándar determinado, pudiendo ser el largo de fila o el tiempo de espera, en ambos casos su valor promedio o máximo. Este estándar puede además mantenerse uniforme a lo largo de toda la red.
- **Disminución del número de solicitudes de modificación.** Al tener una programación más robusta, se aminora la necesidad de hacerle modificaciones. Las solicitudes de modificación presentan complicaciones tales como:
 - Procesar solicitudes en distintos formatos.
 - Analizar la información relevante para evaluar cada solicitud.
 - Difícil reacción dentro de un plazo apropiado.

- **Disminución de fallas humanas.** Al automatizarse el proceso se suprime los errores en el flujo de información que se pueden producir al especificar manualmente horarios, mesaninas, fechas o tipos de servicio de caja.
- **Evaluación de programaciones.** Adicionalmente a su funcionalidad principal, el sistema permite simular programaciones arbitrarias para así apoyar al monitoreo de mesaninas y, de esta manera, contribuir a la mejora continua de la programación.
- **Estandarización de informes y almacenamiento de programaciones.** El uso del sistema permite la generación de una serie de reportes estándares para el análisis y monitoreo del funcionamiento del sistema. Adicionalmente, da la posibilidad almacenar y reutilizar programaciones previas.

4.3.2. Operatividad

Con el sistema desarrollado se pueden generar programaciones que consideren numerosas condiciones. Esto ha permitido incluir requerimientos que si bien no están contemplados en los contratos de subcontratación de personal facilitan las operaciones del *Canal de Ventas* y su control. Las condiciones no contractuales que han sido incorporadas son las siguientes:

- Sólo se habilita la operación de rompefilas cuando la boletería está llena.
- Se impide la formación de las llamadas *ventanas* en la programación. Es decir, no se permite que número de servicios de caja programados aumente (o disminuya) por un intervalo de tiempo menor a una hora.

5. Conclusiones

En este documento se describe la implementación y la puesta en funcionamiento de un sistema computacional para la determinación de la dotación de personal de caja para las boleterías del Metro de Santiago.

El sistema descrito se basa en la combinación de dos modelos que interactúan iterativamente: un modelo de optimización, que determina de manera aproximada la dotación necesaria del personal y, un modelo de simulación, que describe el comportamiento del sistema dada una programación. Lo anterior, incorporando las complejidades del sistema.

El sistema propuesto fue testeado en una prueba piloto en un conjunto reducido de estaciones durante el año 2006. Los resultados obtenidos muestran

beneficios o mejoras en al menos tres planos: el plano económico, organizacional y de calidad de servicio.

El sistema computacional presentado no sólo emplea los modelos sofisticados desarrollados en el proyecto para Metro S.A. Además, ha sido implementado exitosamente en los sistemas informáticos de la compañía mostrando resultados muy satisfactorios.

Como se ha analizado en la Sección 2.1, hay un enorme potencial para el enfoque presentado en este artículo. La metodología y el sistema descrito son aplicables a cualquier problemática de negocio que se caracterice por la definición de los requerimientos de personal para un servicio de demanda muy variable en el transcurso del día. Algunas potenciales aplicaciones son la programación de personal en *call centers*, en las cajas de bancos y supermercados, y en la optimización de los turnos de hospitales.

Agradecimientos: Al Instituto Científico Milenio “Sistemas Complejos de Ingeniería” P04-066-F por el apoyo económico para la concreción de este proyecto.

Referencias

- [1] S. Agnihothri and P. Taylor. Staffing a centralized appointment scheduling department in Lourdes Hospital. *Interfaces*, 21(5):1–11, 1991.
- [2] D. Ashley. A spreadsheet optimization system for library staff scheduling. *Computers & Operations Research*, 22(6):615–624, 1995.
- [3] J. Atlason, M. Epelman, and S. Henderson. Optimizing call center staffing using simulation and analytic center cutting-plane methods. *Management Science*, 54(2):295–309, 2008.
- [4] F. Azadivar. Simulation optimization methodologies. *Proceedings of the Winter Simulation Conference 1999*, 1:93–100, 1999.
- [5] A. Bush. Iterative optimization and simulation of barge traffic on an inland waterway, 2003. Proceedings of the Winter Simulation Conference 2003 2:1751-1756 vol.2.
- [6] M. Byrne and M. A. Bakir. Production planning using a hybrid simulation analytical approach. *International Journal of Production Economics*, 59(1-3):305–311, 1999.
- [7] M. Centeno. A simulation-ILP based tool for scheduling ER staff. *Proceedings of the Winter Simulation Conference 2003*, 2:1930–1938, 2003.

- [8] M. T. Cezik and P. L'Ecuyer. Staffing multiskill call centers via linear programming and simulation. *Management Science*, 54(2), 2008.
- [9] W. Cheung, L. C. Leung, and P. C. F. Tam. An intelligent decision support system for service network planning. *Decision Support Systems*, 39(3):415–428, 2005.
- [10] G. B. Dantzig. A comment on Edie's "Traffic delays at toll booths". *Journal of the Operations Research Society of America*, 2(3):339–341, 1954.
- [11] B. T. Denton, A. S. Rahman, H. Nelson, and A. C. Bailey. Simulation of a multiple operating room surgical suite. *Proceedings of the Winter Simulation Conference 2006*, pages 414–424, 2006.
- [12] L. Edie. Traffic delays at toll booths. *Journal of the Operations Research Society of America*, 2(2):107–138, 1954.
- [13] A. Ernst, H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. An annotated bibliography of personnel scheduling and rostering. *Annals of Operations Research*, 127(1-4):21–144, 2004.
- [14] A. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153(1):3–27, 2004.
- [15] D. Hammond and S. Mahesh. A simulation and analysis of bank teller manning. *Proceedings of the Winter Simulation Conference 1995*, pages 1077–1080, 1995.
- [16] S. G. Henderson. Rostering by iterating integer programming and simulation. *Proceedings of the Winter Simulation Conference 1998*, 1:677–683, 1998.
- [17] Y. Hung and R. C. Leachman. A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations. *Semiconductor Manufacturing, IEEE Transactions*, 9:257–269, 2004.
- [18] E. Ignall, P. Kolesar, and W. Walker. Using simulation to develop and validate analytic models: some case studies. *Operations Research*, 26(2):237–253, 1978.
- [19] B. Kim and S. Kim. Extended model for a hybrid production planning approach. *International Journal of Production Economics*, 73(2):165–173, 2001.

- [20] A. Mason, D. Ryan, and D. Panton. Integrated simulation, heuristic and optimisation approaches to staff scheduling. *Operations Research*, 46(2):161–175, 1998.
 - [21] Y. Naveh, Y. Richter, Y. Altshuler, D. Gresh, and D. P. Connors. Workforce optimization: Identification and assignment of professional workers using constraint programming. *IBM Journal of Research and Development*, 51(3/4):263–279, 2007.
 - [22] X. Qi and J. F. Bard. Generating labor requirements and rosters for mail handlers using simulation and optimization. *Computers & Operations Research*, 33(9):2645–2666, 2006.
-

6. Anexo

Modelo de Programación Lineal Entera

A continuación se detalla el modelo de programación entera mixta utilizado como base del *Módulo de Asignación* del sistema. Como hemos indicado, para cada mesanina–día se resuelve una instancia diferente. Antes de presentar el modelo describimos la notación y parámetros necesarios.

La jornada de operación del servicio está dividido en Q *bloques* de igual duración $q = 1, \dots, Q$. Estos bloques constituyen la unidad utilizada para discretizar el tiempo en el planteamiento del modelo. En particular se utilizaron bloques de 15 minutos, por lo que una jornada tipo de 6:30 a 22:30 horas, es dividida en 64 bloques.

Las contrataciones deben ser realizadas por un tipo específico de operario. Para el problema estudiado, el tipo de operario depende tanto del equipo que utiliza (si ocupa POS o no) y de la posición de trabajo en la mesanina (en boletería –*puesto fijo*– o como rompe-filas –*puesto móvil*–). De esta manera, se cuenta con 4 tipos de operarios: “Fijo con POS”, “Fijo sin POS”, “Móvil con POS” y “Móvil sin POS”. El conjunto de estos cuatro tipos de puestos de trabajo se denota por P .

Las contrataciones deben ser realizadas en turnos de duración acorde a lo estipulado en los contratos. Para fines del modelo, definimos L como el conjunto de las duraciones de los turnos, medidas en número de bloques. En el caso particular considerado, los turnos pueden durar 2, 4 o 16 horas. Como los bloques son de 15 minutos, esto significa que $L = \{8, 16, 64\}$.

Los cajeros deben atender diferentes tipos de transacciones. El conjunto de tipo de transacciones a considerar se denota por T . En la aplicación se consideraron dos tipos de transacciones: “compra de boleto” y “carga de

tarjeta”.

Para el planteamiento de ciertas restricciones es necesario considerar sólo algunos tipos de operarios. Para esto, definimos los siguientes subconjuntos de P:

- Puestos fijos: $PF = \{ \text{"Fijo con POS"}, \text{"Fijo sin POS"} \}$.
- Puestos móviles: $PM = \{ \text{"Móvil con POS"}, \text{"Móvil sin POS"} \}$.
- Puestos con POS: $PC = \{ \text{"Fijo con POS"}, \text{"Móvil con POS"} \}$.
- Puestos sin POS: $PS = \{ \text{"Fijo sin POS"}, \text{"Móvil sin POS"} \}$.
- Puestos que pueden atender una transacción tipo t : $P(t)$.

Las Tablas 3 y 4 muestran los otros parámetros necesarios y los valores utilizados. Los *parámetros generales* corresponden a los datos que no varían entre mesaninas o días diferentes; los parámetros *por mesanina* son valores que dependen de la instancia particular a resolver.

Parámetro	Notación	Descripción
Eficiencia	EF	Fracción del tiempo que un servidor está efectivamente disponible.
Tiempo de Atención	A_p^t	Tiempo medio de atención de una transacción de tipo t por un puesto tipo p , en segundos.
Costo de contratación	C_{pl}	Costo de contratación de un puesto tipo p en un turno de largo l .
	α	Parámetro de control, relación mínima entre tasa de llegada de transacciones y tasa nominal de atención.
Largo de bloque	LQ	Duración de un bloque, en segundos.

Cuadro 3: Parámetros generales.

Adicionalmente, se definen el conjunto $R \subseteq \{1, \dots, Q\}$ de períodos en los que hay que corregir la asignación previa y número mínimo de servidores a programar en el período r de acuerdo a la corrección, g_r para cada período $r \in R$. Inicialmente el conjunto R es vacío y se va actualizando en cada iteración entre los módulos.

Para representar las decisiones de contratación se define un conjunto de variables enteras no negativas, N_{pql} que representan la cantidad de puestos de venta del tipo p que comienzan a operar en el período q por un bloque de largo l , para cada combinación posible. Adicionalmente, se definen los siguientes grupos de variables auxiliares:

Parámetro	Notación	Descripción
Demanda	D_q^t	Número de transacciones de tipo t en el bloque q .
Dotación mínima	M^{\min}	Número mínimo de servidores que deben estar trabajando en todo momento. Puede ser nulo.
Dotación máxima por tipo	M_p^{\max}	Número máximo de puestos de tipo p que pueden estar trabajando simultáneamente.
Dotación máxima puestos fijos	MaxFijos	Número máximo de puestos fijos que pueden estar trabajando simultáneamente, es decir $\sum_{p \in PF} M_p^{\max}$.
	f_{qt}	Representa la fracción de la carga de trabajo del bloque q que corresponde a transacciones de tipo t . En la implementación se definió como $(D(q, t) / \sum_{t' \in T} D(q, t'))(1/A(q, t))$.

Cuadro 4: Parámetros por mesanina

- para cada tipo de operario p , transacción t y bloque horario q , la variable entera no negativa Y_{pq}^t representa el número de transacciones de tipo t atendidas por los puestos tipo p en q ;
- para cada par de tipo de operario-bloque horario (p, q) , se definen dos variables. La primera es una variable entera, X_{pq} que representa el número de servidores de tipo p activos en el bloque q , y la segunda es una variable binaria w_{pq} , la que toma el valor 1 si en el bloque q se inicia algún servidor tipo p ;
- para cada bloque q se definen dos conjuntos de variables binarias que indican si en el lapso de media hora (2 bloques) hubo modificaciones de dotación y que son utilizadas para evitar la formación de *ventanas*. La variable B_q toma un valor igual a 1 si la cantidad de servidores activos en q es menor a la cantidad de servidores activos en el período $q - 2$ y, toma un valor igual a 0 en caso contrario. La segunda variable llamada S_q toma un valor igual a 1 si la cantidad de servidores activos en $q + 2$ es mayor a la cantidad de servidores activos en el período q y, toma un valor igual a 0 en caso contrario.

Con esta notación, el modelo implementado es el que se muestra en la página siguiente. El primer conjunto de restricciones liga las variables N_{pql} de inicio de servicio con las X_{pq} que cuentan el número de servidores activos en cada momento. Las restricciones siguientes garantizan la cantidad mínima de cajeros activos necesaria y que no se programen más servidores que la capacidad de la mesanina.

El cuarto conjunto de restricciones garantiza que no se habiliten puestos móviles si no están en operación todos los puestos fijos. Esta restricción es de

interés operativo de la empresa y se fundamenta en que la operación de los puestos fijos es más eficiente y segura.

La quinta restricción relaciona la demanda en número de transacciones (lado izquierdo, normalizado por el parámetro α) y la capacidad nominal de atención de los cajeros en operación en el bloque q .

La sexta restricción incorpora las correcciones a hacer a la programación que han sido detectadas por el módulo de simulación.

Las últimas cuatro restricciones se incluyen para evitar la formación de las llamadas *ventanas temporales*. Por motivos de control de operación es importante que no se produzcan variaciones abruptas en la dotación en un período de tiempo muy corto. Específicamente en este caso, se está impidiendo que la dotación aumente en un determinado instante y luego vuelva a reducirse a la media hora o que se reduzca y luego vuelva a crecer en media hora.

La función objetivo utilizada corresponde a minimizar los costos totales de contratación del personal de caja.

Modelo 1: Programación de cajeros

$$\text{minimizar} \quad z = \sum_{p \in P} \sum_{l \in L} \sum_{q=1}^Q C_{pl} N_{pql}$$

sujeto a

$$\begin{aligned}
X_{pq} &= \sum_{l \in L} \sum_{r=q-l+1}^q N_{prl} && \text{para todo } p \in P, q = 1, \dots, Q, \\
X_{pq} &\leq M_{pq}^{\text{Max}} w_{pq} && \text{para todo } p \in P, q = 1, \dots, Q, \\
\sum_{p \in P} X_{pq} &\geq M^{\text{min}} && \text{para todo } q = 1, \dots, Q, \\
\sum_{s \in \text{PF}} X_{sq} &\leq \text{MaxFijos} \cdot w_{pq} && \text{para todo } p \in \text{PM}, q = 1, \dots, Q, \\
\alpha \sum_{t \in T} D_q^t &\leq LQ \cdot \sum_{t \in T} \sum_{p \in P(t)} f_{qt} X_{pq} && \text{para todo } q = 1, \dots, Q, \\
\sum_{p \in P} X_{pr} &\geq g_r && \text{para todo } r \in R, \\
\sum_{p \in P} X_{pq} - \sum_{p \in P} X_{p,q+2} &\leq B_q \sum_{p \in P} M_{pq}^{\text{Max}} && \text{para todo } q = 1, \dots, Q, \\
\sum_{p \in P} X_{p,q+2} - \sum_{p \in P} X_{pq} &\leq S_q \sum_{p \in P} M_{pq}^{\text{Max}} && \text{para todo } q = 1, \dots, Q, \\
B_q + S_q &\leq 1 && \text{para todo } q = 1, \dots, Q, \\
B_{q+1} + S_{q-1} &\leq 1 && \text{para todo } q = 1, \dots, Q, \\
N_{pql} &\in \mathbb{Z}_+ && \text{para todo } p \in P, l \in L, q = 1, \dots, Q, \\
X_{pq} &\in \mathbb{Z}_+ && \text{para todo } p \in P, q = 1, \dots, Q, \\
w_{pq} &\in \{0, 1\} && \text{para todo } p \in P, q = 1, \dots, Q, \\
B_q, S_q &\in \{0, 1\} && \text{para todo } q = 1, \dots, Q.
\end{aligned}$$

MEJORAMIENTO DE LA GESTIÓN Y USO DE TICS EN LAS MIPYMES A TRAVÉS DE SOFTWARE LIBRE

ÁNGEL JIMÉNEZ MOLINA*
JUAN D. VELÁSQUEZ SILVA*
ALEJANDRO FUENTES DE LA HOZ*

Resumen

Las tecnologías de información y comunicaciones (TICs) han permitido un mejoramiento de los procesos de negocios, labores administrativas y la gestión en general de diversas organizaciones. Sin embargo, por razones de costos, en la mayoría de los casos las TICs no son utilizadas por las Mipymes. En la última década ha surgido el paradigma tecnológico conocido como software libre. Las Mipymes enfrentan un rezago tecnológico en cuanto al apoyo con TICs, principalmente por el costo del software y desconocimiento de la existencia y eficiencia del software libre. En este paper se prueba para los sectores industriales agrícola/frutícola, acuícola y maderero, por medio de una encuesta, que tal rezago tecnológico es un hecho, y que es posible solucionarlo mediante el uso de software libre. Luego se indagan los requerimientos tecnológicos funcionales para la incorporación de TICs en las Mipymes, a través del estudio de sus procesos de negocios claves. Finalmente, por medio del análisis de la oferta de soluciones basadas en software libre, se seleccionan herramientas que satisfagan tales requerimientos. Se concluye que estas herramientas pueden ser utilizadas como base para adaptar soluciones concretas por medio de proyectos de implantación, pues se advierte que no para todos los requerimientos existe software libre que los satisfaga directamente.

Palabras Clave: Software libre, Mipymes, Procesos de Negocios.

*Departamento de Ingeniería Industrial, Universidad de Chile

1. Introducción

Las tecnologías de información y comunicaciones (TICs) han sido ampliamente usadas en instituciones de diversa índole, para mejorar sus procesos de negocios, labores administrativas y la gestión en general [2]. Sin embargo, por razones de costos, en la mayoría de los casos las TICs son sólo para un segmento muy particular de organizaciones, es decir, para aquellas que pueden solventar las fuertes inversiones que demanda la adquisición, implantación, mantención o incluso el desarrollo de una solución tecnológica.

Por lo anterior, pareciera ser casi imposible que una micro, pequeña o mediana empresa (Mipyme) pudiese acceder a las bondades de las TICs. Sin embargo, en la última década se ha producido un movimiento en el ámbito del desarrollo de software, que podría transformarse en una solución real para las Mipymes de cara a la adopción de TICs en su devenir diario. Se trata del software libre (open source), el cual puede ser definido como cualquier programa cuyo código fuente fue escrito con la intención declarada de dejarlo disponible para ser usado, y eventualmente modificado, por terceros. El software libre es usualmente desarrollado como una colaboración pública entre varios programadores y puesto a disposición de la comunidad en forma gratuita. Sin embargo, el desconocimiento de la existencia y eficiencia del software libre es una de las barreras que limitan su uso por las Mipymes.

Las Mipymes enfrentan un rezago tecnológico en cuanto al apoyo con TICs de los procesos de negocios claves de su cadena de valor. Esto sin duda afecta su efectividad operacional y capacidad para generar y sostener ventajas competitivas en el tiempo [7]. En este paper se prueba que tal rezago tecnológico es un hecho, y que el software libre es una solución apropiada para algunas de las necesidades. Esto se efectúa por medio de la investigación del estado actual en el que se encuentra el uso de TICs en tales Mipymes, a través de la aplicación de una encuesta. Además, se indagan por medio del estudio de sus procesos de negocios claves, los requerimientos tecnológicos específicos que estas empresas tienen para incorporar TICs, principalmente soluciones de infraestructura tecnológica y de negocios. Finalmente se analiza la oferta de soluciones dispuestas por el software libre, y se evalúa la pertinencia de su uso en cada requerimiento.

La investigación se focalizó en tres sectores industriales: (1) agrícola / frutícola, (2) acuícola y (3) maderero, los cuales agrupan a Mipymes nacionales con un interesante potencial de crecimiento asociativo, incluso dado por ventajas comparativas intrínsecas.

La organización de este paper considera en primer lugar el trabajo rela-

cionado sobre el uso de software libre en las Mipymes nacionales, por medio de la sección 2. A continuación, la sección 3 indaga el estado del uso de TICs en las Mipymes de los sectores indicados, por medio de la exposición de los resultados de una encuesta que se aplicó sobre una muestra de estas empresas. La sección 4 efectúa un análisis de los procesos de negocios claves de la cadena de valor de las Mipymes. A partir del análisis de procesos, la sección 5 determina los requerimientos tecnológicos funcionales para las Mipymes, los cuales son usados en la sección 6 para seleccionar herramientas basadas en software libre que los satisfagan. El paper finaliza con las conclusiones, donde además se enuncia el trabajo futuro requerido para la continuidad de esta investigación.

2. Trabajo Relacionado

Según el mejor conocimiento de los autores, no se registran estudios nacionales pasados o recientes orientados a evaluar necesidades específicas de TICs por las Mipymes, que se puedan satisfacer con software libre. Uno de los estudios que entrega información de interés se documenta en [20]. Sin embargo, a pesar de lo interesante de los datos recopilados, es en el ámbito particular del Gobierno, lo que en general no permite hacer algún tipo de conclusión respecto a nuestro país ni en particular respecto a las Mipymes.

3. Estado del Uso de TICs en las Mipymes Nacionales

Por cada Mipyme se generó un documento de encuesta que posteriormente fue analizado individual y agregadamente. El objetivo de la encuesta es establecer un catastro de las TICs en uso, ya sea software libre o no, y por qué. Al mismo tiempo, identificar la percepción, prejuicios o motivaciones que estas empresas pudieran tener sobre el uso de las TICs. Se indagaron además los aspectos que han dificultado o imposibilitado la implementación de las TICs, sus experiencias tanto exitosas como aquellas que no han implicado mejoras significativas en la productividad o eficiencia de sus procesos.

3.1. Muestreo y Caracterización de Datos

Se consideró una muestra de Mipymes distribuidas en las regiones del país en las que se concentra la actividad de los sectores acuícola, maderero y agríco-

la/frutícola¹ . La muestra se constituyó utilizando las siguientes fuentes de información: directorio industrial de Chile [15], Chile alimentos [16], directorio forestal Lignum [17] y el directorio de acuicultura y pesca [18]. Del total de Mipymes contactadas, se encuestaron 37, cifra compuesta por un 29,73 % del sector agrícola/frutícola, 32,43 % del sector acuícola y el resto del sector maderero. Complementariamente a la aplicación en terreno de la encuesta, se recaudó por medio de observación, información adicional que permitiera el posterior modelamiento del negocio y sus procesos.

3.2. Caracterización de la Encuesta

Las líneas de indagación de la encuesta contemplaron en principio la obtención de información general de la Mipyme en los ámbitos de su propiedad, rubro específico, intensidad de mano de obra, facturación, antigüedad, etc. En segundo lugar, la identificación de los procesos de negocios esenciales de su cadena de valor, por ejemplo en el ámbito de la relación con los clientes, los proveedores, y sus actividades operacionales y transaccionales claves. En tercer lugar, la encuesta consideró la obtención de datos relativos a la percepción y aptitudes hacia el uso de las TICs en los procesos de negocios anteriores. De esta forma, complementándose con observación en terreno, se conformó información de base para diagnosticar la situación actual del uso de las TICs en los procesos de negocios de su cadena de valor. En cuarto lugar, se indagó la disponibilidad y tipo de equipamiento informático, en caso de existir.

3.3. Principales Resultados de la Encuesta

Esta sección resume los principales resultados de la encuesta. En primer lugar se muestra la situación actual de la plataforma informática de base de las Mipymes, para luego identificar la tecnología utilizada para el soporte de sus principales procesos de gestión. Finalmente se muestran resultados sobre la disposición que en general las Mipymes tienen para la implantación de TICs.

3.3.1. Situación Actual de la Tecnología de Base de las Mipymes.

Esta sección analiza la plataforma informática con que cuentan las Mipymes². El 94,54 % de las empresas encuestadas poseen computadores. En aquellas que no poseen computadores, se argumentaron razones de precio de

¹Diversas ciudades fueron consideradas en el estudio, tales como: Chillán, Puerto Montt, Rengo, Temuco, Romeral, Buin, Ovalle, Loncoche, Rancagua, Chimbarongo, Los Andes, Santiago (Peñaflor, Peñalolén, Vitacura, La Granja), Talca, Tongoy, La Serena, entre otras.

²De las empresas encuestadas, solamente 2 no poseían ningún tipo de equipamiento informático, pertenecientes al rubro maderero. Consultadas de por qué no poseían equipamiento informático, las razones tienen relación con el precio de adquisición y que no visualizan su utilidad en el corto plazo.

adquisición, mantención y valor del software necesario a instalar. En cuanto a software, en su gran mayoría se trata de paquetes estándares, tales como sistemas operativos, procesadores de texto, etc., de los cuales no se tiene información de cómo fueron adquiridos. Adicionalmente, el 45,71 % de las Mipymes encuestadas posee red local. Los sectores que más empresas tienen sin esta infraestructura son el acuícola y el maderero. En cambio, en el agrícola/frutícola, las empresas que cuentan con red superan en número a las que no tienen. Con respecto al acceso a Internet, se comprueba que el 77,14 % de las empresas de la muestra cuentan con enlace. Sin embargo sólo el 56 % de éstas últimas - 43 % del total de la muestra - posee red de área local. Esto indica que en una primera etapa las empresas invierten en acceso a Internet y luego, invierten en la red local - en promedio 1 de cada 2 empresas con enlace posee red local -. Lo anterior se puede deber a que existe un sólo computador conectado a Internet - el del jefe o gerente -, fundamentalmente, como pudo detectarse, para labores de correo electrónico, acceso a bancos, sitios del Gobierno, etc.

La red local se considera en una segunda etapa y no está claro si los usuarios de éstas también tendrán acceso a Internet. Lo anterior resulta gravitante a la hora de evaluar cual sería el modelo de operación más adecuado de las soluciones de negocios. Por ejemplo, el no contar con acceso a Internet por parte de todos los computadores conectados a una red local, constituye un elemento que dificulta la utilización de soluciones en formato de Application Service Provider (ASP), entre otros.

El tipo de conexión más utilizado es el ADSL -50 %-, seguida por el Cable. Otra característica a destacar es que las empresas pesqueras muestran la mayor diversidad de conexiones, debido a la distribución geográfica y a que requieren infraestructura inalámbrica para las operaciones sobre el mar o lagos. La existencia mayoritaria de conexiones a Internet vía ADSL y Cable, garantiza un buen ancho de banda para la mayoría de las aplicaciones de software libre que se pretenden implementar en un proyecto de esa naturaleza, lo cual potencia aún más su factibilidad.

Un 91,43 % de las empresas que cuentan con Internet, utilizan el correo electrónico. El único caso en el que se detectó presencia de Internet, pero sin uso de correo electrónico, fue en el sector maderero, pero aisladamente. En la mayoría de los casos la infraestructura tecnológica de hardware no sería un gran obstáculo para la implantación de un proyecto de Software Libre, por cuanto existe una arquitectura básica sobre la cual operar.

3.3.2. Soporte Tecnológico de los Procesos de Gestión de las Mipymes.

Se indago la frecuencia con que las Mipymes efectúan actividades relacionadas a la gestión de sus clientes. Esta información de frecuencia contempla,

entre otras actividades, la actualizan de los datos de sus clientes, la captación de pedidos y requerimientos, la gestión de venta y postventa, la segmentación de sus clientes, etc. Se constató que cerca del 51 % de las Mipymes efectúa gestión de sus clientes en períodos muy largos de tiempo -una vez al año o mensualmente-, y sólo un 24 % lo hace diariamente. Los encuestados indicaron como razones de su inadecuada gestión de sus clientes, el no contar con herramientas tecnológicas habilitadoras y, en otros casos, no poseer fundamentación teórico-práctica relativa a la gestión de clientes.

		Ejecución interna			Ejecución con apoyo de gestor externo		
		Automatizado	Manual con ayuda de PC	Manual	Automatizado	Manual con ayuda de PC	Manual
	Gestión de clientes	43,24 %	43,24 %	13,51 %	0 %	0 %	0 %
	Producción	35,14 %	45,95 %	16,22 %	0 %	2,7 %	0 %
	Contabilidad	35,14 %	5,41 %	5,41 %	27,03 %	27,03 %	0 %
	Renuméricaciones	40,54 %	35,14 %	16,22 %	2,7 %	5,41 %	0 %

Tabla 1. Modalidad y Nivel de Apoyo Tecnológico para la Ejecución de Procesos Claves de las Mipymes.

Como se observa en la tabla 1, todos los procesos claves de las Mipymes se ejecutan mayoritariamente de forma interna, excepto, como es predecible, el proceso de contabilidad, que en un 54,6 % se hace por medio del apoyo de un gestor externo. Sin embargo, en los gestores externos no se aprecia un uso de herramientas de software superior al 50 % de los casos. De verificararse apoyo externo para los procesos claves restantes, su uso de software es insignificante.

En cuanto a la ejecución interna, un gran porcentaje de las Mipymes aún utiliza procesos manuales, excepto para la contabilidad. En efecto, en casi todos los procesos claves esta cifra es superior al 50 %, alarmando la ejecución de los procesos de producción, que en un 62,17 % es meramente manual. Sin embargo, al menos se aprecia un grupo mayoritario dentro de las Mipymes con ejecución manual que ha incorporado el uso básico del PC, lo cual, sin embargo, no puede ser considerado sinónimo de automatización. En su gran mayoría se trata de la utilización de herramientas de ofimática - planillas de cálculo, procesadores de texto -, pero en ningún caso, herramientas de alto nivel - ERP, Workflow, CRM, almacenamiento de datos, etc. -, para mejorar significativamente la eficiencia y eficacia de los procesos.

Este análisis confirma el problema de rezago tecnológico de las Mipymes que recurrentemente se observa en la realidad. Por lo tanto existe potencial para apoyar con software los procesos claves de las Mipymes que en estos momentos se realizan manualmente y, en casos aislados, con algún nivel de automatización. La página Web de productos, la gestión del inventario y pérdidas, más la gestión de la producción, serían las áreas donde el uso del software tendría un mayor impacto según las Mipymes.

A nivel desagregado por sector, el resumen de la tabla 2 muestra la misma conclusión, donde se han ennegrecido las mayores necesidades percibidas por

las Mipymes. Claramente estas corresponden a las mismas enunciadas previamente a un nivel agregado. Destaca sin embargo el sector acuícola, que además manifiesta inquietudes no despreciables en los procesos de gestión de los proveedores, de la distribución y de los clientes.

	Sector acuícola	Sector agrícola / frutícola	Sector maderero
Proceso o área			
Página Web de productos	16,7 %	18,2 %	33,3 %
	0,0 %	9,1 %	4,2 %
	33,3 %	27,3 %	25,0 %
	11,1 %	9,1 %	4,2 %
	0,0 %	0,0 %	4,2 %
	11,1 %	0,0 %	0,0 %
	5,6 %	18,2 %	16,7 %
	0,0 %	0,0 %	4,2 %
	5,6 %	9,1 %	4,2 %
	0,0 %	9,1 %	0,0 %
	5,6 %	0,0 %	4,2 %
	11,1 %	0,0 %	0,0 %

Tabla 2. Desagregación por Sector de la Autopercepción del Rezago.

Con un nivel de desagregación más, dado por los segmentos de empresas - micro, pequeña y mediana -, se aprecia similarmente la misma conclusión. En efecto, la autopercepción de necesidades tecnológicas prioritarias de las Mipymes se encuentran dadas por el apoyo a los procesos de gestión de inventarios, gestión de clientes, gestión de la cobranza y página Web de productos. Además, en dos de los sectores (agrícola/frutícola y maderero), las microempresas consideran prioritario la página Web de productos.

Por otro lado, en los sectores madereros y acuícola las pequeñas empresas coinciden en la prioridad para la gestión de inventarios y pérdidas. Sobre las medianas empresas nada se puede afirmar claramente.

3.3.3. Disponibilidad Organizacional Hacia la Implantación de TICs.

En cuanto al nivel de conocimiento del software libre y su penetración en las empresas encuestadas un 64,86 % de las Mipymes de la muestra declaró haber escuchado sobre éste. De las empresas que conocen el software libre, sólo un 46,1 % entiende la diferencia con respecto a una aplicación propietaria. Esto muestra que la ausencia de soluciones de software libre en más Mipymes se debe en su gran mayoría al desconocimiento de su existencia y no así a razones de otra índole, como pueden ser dificultad de adquisición, instalación, mantención, credibilidad, etc.

A pesar del desconocimiento de lo que es el software libre, el 91,89 % mues-

tra interés por recibir información respecto a este tema y analizar seriamente su posibilidad de implantación en sus respectivas instituciones. Del 64,9 % de los encuestados que no conoce la existencia del software libre, desearía recibir información, mientras que del total de encuestados, sin importar el rubro al que pertenecen más del 90 % se muestra interesado en las bondades de este tipo de soluciones.

En cuanto a la aceptación que tiene dentro de las Mipymes la adopción de TICs en la mejora de sus procesos de negocios ejecutantes y de gestión, la mayoría está de acuerdo en que su incorporación ayudará a mejorar los procesos productivos y de gestión. Son las empresas relacionadas con el sector acuícola, las que estarían más proclives y entusiasmadas a la adopción de TICs en sus procesos. En efecto, durante las visitas a terreno se manifestaron de acuerdo en apoyar el desarrollo de un proyecto de implantación de herramientas de software libre, actuando como usuarios proactivos en el uso de las aplicaciones involucradas en el proyecto. Si bien todos los sectores consideran valiosas las TICs, no hay consenso respecto a la complejidad de su implantación.

Algunos sectores como el agrícola/frutícola, consideran sencilla la implantación de un proyecto TIC, aunque en las visitas en terreno no se observó una real inversión en este ámbito.

Otro dato interesante es que el 83,78 % de las Mipymes encuestadas, considera que su empresa sería más competitiva al implantar TICs. Se concluye que las TICs son percibidas como fuentes verdaderas de mejoramiento de la competitividad dentro de la empresa. Por otro lado, es interesante notar que más del 50 % de los encuestados, se declara de acuerdo con implementar un proyecto que involucre el uso de herramientas de software libre y participaría en su desarrollo.

Para poder priorizar cuáles serían las características más relevantes, desde punto de vista de las Mipymes encuestadas, para la implantación de un proyecto de software libre, se les solicitó que priorizaran un conjunto de características deseables. Sin duda, la característica que más relevancia tiene para las Mipymes, es la sencillez en el uso de la herramienta, sugiriendo la importancia que tendría la capacitación dentro de un proyecto de implantación de software libre. La segunda característica relevante corresponde a la adaptabilidad de la herramienta a los requerimientos puntuales de las empresas. Finalmente, las Mipymes no consideran valioso que la herramienta sea estándar genérico, lo cual muestra que aún no logran percibir las facilidades que otorgan las soluciones de TICs con estas características para la comunicación con otras plataformas.

4. Análisis del Negocio y sus Procesos.

El objetivo de esta sección es caracterizar genéricamente los procesos de negocios de los sectores analizados, para definir en la siguiente sección qué tipo de herramientas tecnológicas se requieren para las Mipymes, es decir, sus requerimientos tecnológicos funcionales. Se ha escogido el enfoque de procesos de negocios como la vía de elicitation de requerimientos de apoyo tecnológico, pues estos permiten una alineación estratégica de las aplicaciones, en el entendido que los procesos operacionalizan el modelo de negocios. Por lo tanto, representa una oportunidad de aplicación a la realidad de las Mipymes nacionales, cuyos procesos de negocios son en general simples y con un gran rezago de apoyo tecnológico como se concluyó en la sección anterior.

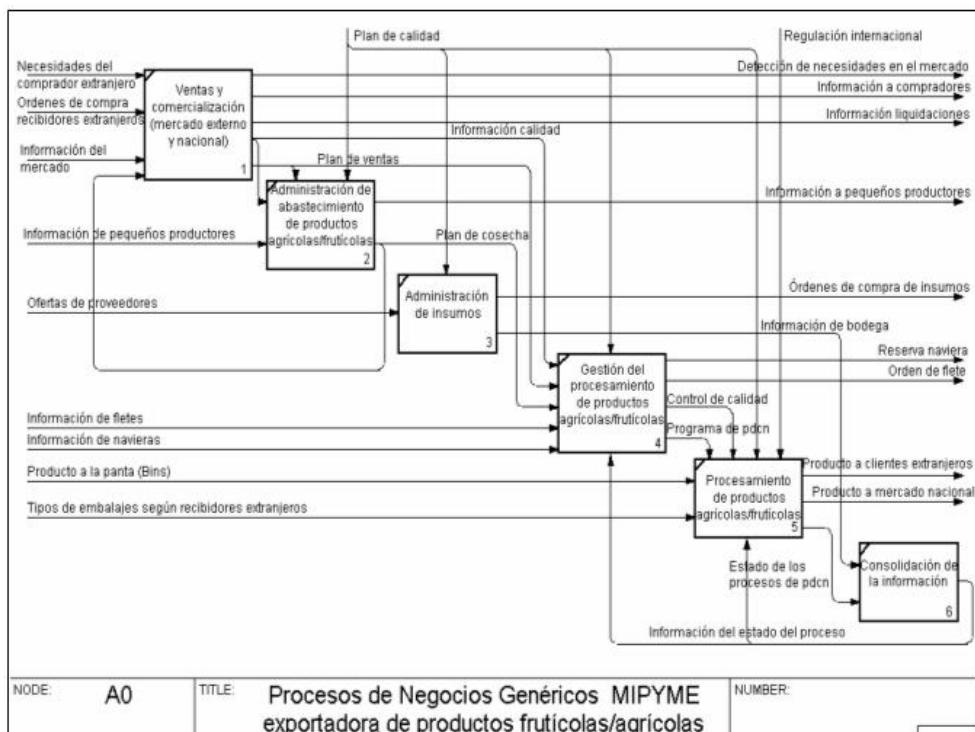


Figura 1. Primer Nivel de Procesos Genéricos para Mipyme exportadora de productos frutícolas/agrícolas.

4.1. Sector Agrícola/Frutícola

Un ejemplo claro de modelos de negocios exitosos son las empresas medianas exportadores de productos agrícolas/frutícolas. Respecto a éstas últimas

mas, se pudo apreciar que no necesariamente cuentan con cultivos, sino que adquieren el producto de pequeños productores agrícolas. Esto arrastra positivamente a las Mipymes a adherirse a este modelo de negocios, que principalmente depende de los mercados internacionales, constituyéndose en un interesante sector a ser analizado. Por lo tanto, existe la necesidad de apoyar a este grupo de Mipymes, las cuales presentan un rezago en el apoyo de sus procesos de negocios centrales con TICs y prácticas de gestión.

En primer lugar es necesario tener en cuenta que los pequeños productores planifican los volúmenes de sus plantaciones de una temporada a la otra en base a las estimaciones agregadas de la demanda externa, manteniendo una interacción muy estrecha con las empresas dedicadas al procesamiento del producto, quienes manejan los vínculos con recibidores internacionales.

Por otro lado, los pequeños productores monitorean y mantienen permanentemente sus cultivos con asesoría directa de agrónomos pertenecientes a las empresas procesadoras. Las cosechas, dependiendo del tipo de producto, se efectúan parcializadas durante la temporada, dependiendo del comportamiento de la demanda; sin embargo, en algunos casos la cosecha debe hacerse de una vez, pues compromete la calidad del producto. Este modelo de negocio contempla la entrega de los productos cosechados a las plantas procesadoras, quienes se encargan de efectuar una clasificación de acuerdo a niveles de calidad - especie, variedad, tamaño, color, aspecto en su forma, etc.-

Los modelos de procesos genéricos que aquí se muestran, se desarrollaron a partir del conocimiento del rubro obtenido mediante las encuestas y conversaciones con las Mipymes. Estos modelos son la base para la decisión de las TICs que otorgan más valor a su negocio dadas las necesidades detectadas. El modelamiento se efectuó con la técnica IDEF0 [22].

En concreto, los procesos de negocios genéricos de las Mipymes agrícolas/frutícola exportadoras, pueden describirse en primer lugar considerando sus entradas y salidas globales. En efecto, en la figura 1 se aprecian flujos de entrada tales como “necesidades del comprador extranjero”, “órdenes de compra recibidores extranjeros”, “información de pequeños productores”, “información de fletes”, etc. Adicionalmente, los flujos de salida son aquellos tales como “información a compradores”, “información liquidaciones”, “información a pequeños productores”, “producto a clientes extranjeros”, “producto a mercado nacional”, etc.

Dentro de las actividades del modelo genérico de procesos para este sector, se encuentra “ventas y comercialización”, que corresponden a la gestión comercial realizada con los compradores - recibidores extranjeros -, donde se incluyen las actividades de análisis de mercado, gestión de cartera, negociación de precios y cotizaciones, seguimiento de postventa, generación de información

para liquidaciones³ de productores⁴, generación de plan de venta para operaciones, decisiones de entrega, etc.

En segundo lugar se tiene la “administración de abastecimiento”, punto crucial, ya que corresponde a la administración de la relación con los productores (pequeños agricultores), la estimación de las cosechas para la generación de planes de producción y de ventas, así como la relación técnica con productores que permitan un aseguramiento de calidad en origen.

La “administración de insumos” corresponde a la gestión de compras e inventario de los insumos necesarios para empacar el producto, así como los necesarios por la empresa en general. Depende de información consolidada por frigorífico acerca de las ventas y embarques y se coordina con packing para la entrega de los embalajes.

La “gestión de la producción” corresponde a todas las actividades de programación de operaciones, como son la generación del plan de producción, el plan de empaques, la coordinación de los embarques y otras actividades del ciclo productivo. Aquí se incluye también la gestión de calidad. La ejecución de la producción corresponde a las actividades de producción propias del negocio exportador frutícola/agrícola, es decir, el proceso de transformación del producto para cumplir las especificaciones requeridas por un determinado cliente. Este proceso incluye la recepción del producto, el procesamiento, control de calidad, almacenamiento en frigorífico y despacho.

Se incluye adicionalmente en el diagrama una caja, que no es una actividad, sino que representa el monitoreo permanente del estado global y detallado del proceso para el control y gestión de éste. Se ha denominado “consolidación de la información”.

Es importante acotar que en el informe original de esta investigación [21], se desarrolló además un modelamiento genérico de los procesos de negocios del subsector congelados y deshidratados de fruta”.

4.2. Sector Acuícola

A continuación se muestra una descripción genérica de los modelos de negocios y procesos claves del sector acuícola, los que se generalizaron a partir del conocimiento detallado de dos de las industrias incluidas en la muestra. Estas industrias corresponden a la del ostión y a la salmonera.

Con respecto a la industria salmonera es importante aclarar que se ha tomado como ejemplo por su indiscutible importancia en el sector acuícola. Sin embargo, esta industria no está directamente relacionada con Mipymes,

³Una liquidación corresponde al pago diferido que se hace a un productor por la entrega de su producto a la exportadora (Ver siguiente pie de página).

⁴Se entienden como productores a los agricultores pequeños que proveen de su producto a las empresas exportadoras, para que estas procesen el producto, lo embalen y lo coloquen al mejor precio en los mercados internacionales.

puesto que sus participantes son grandes empresas, muchas de clase mundial. El valor de analizar esta industria se haya en que existe un conjunto no despreciable de Mipymes relacionadas con las grandes salmoneras, ya sea como proveedoras o bien aguas abajo agregando valor al salmón por medio de enlatado o procesamiento básico del producto, muchas de las cuales incluso exportan. Por lo tanto, es necesario comprender el modelo de negocios salmonero y sus procesos como un todo, para lograr una visión de las necesidades tecnológicas de las Mipymes indicadas que, por formar parte de una industria de un nivel competitivo altísimo, deben solucionar cuanto antes su rezago tecnológico.

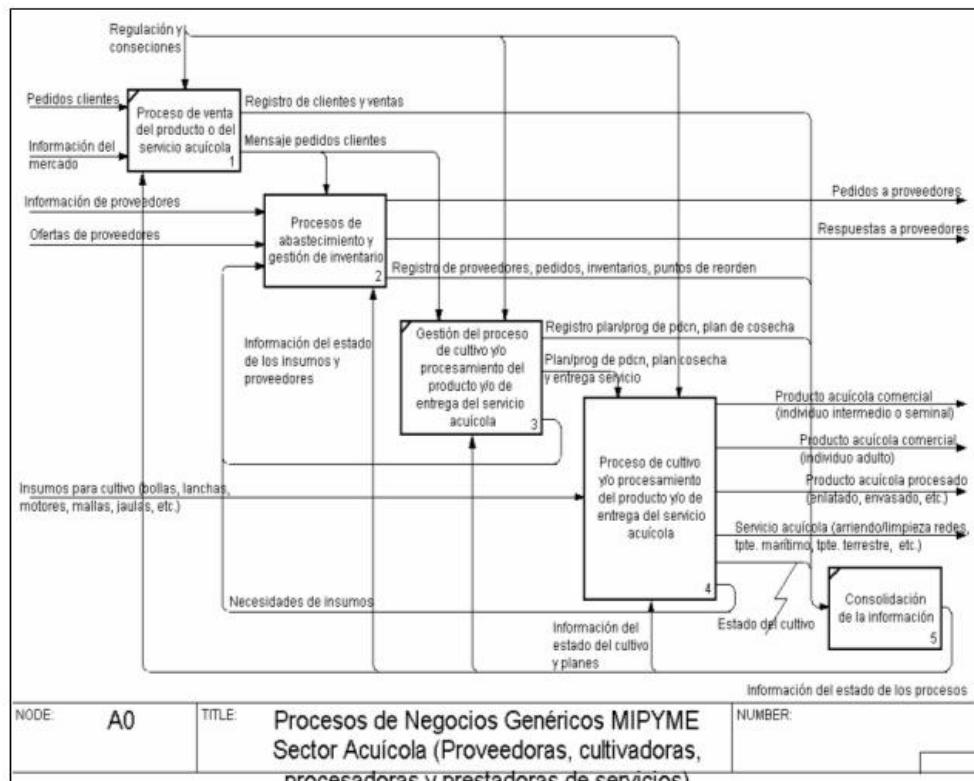


Figura 2. Primer Nivel de Procesos Genéricos para Mipyme Sector Acuícola.

En la industria del cultivo, procesamiento, cosecha y venta del ostión, se identifican dos tipos de actores principales. En primer lugar, se tiene a empresas de envergadura mediana con grandes instalaciones en planta, y por otro, a pequeñas empresas por lo general actúan en el mercado asociadas como gremio. El proceso de cultivo es extensivo, y contempla, en primer lugar, la captación de la semilla del ostión (1 a 2 mm.), su fijación en mallas especiales y la mantención periódica del cultivo suspendido para obtener individuos comerciales (70 a 75 mm.), los cuales se encuentran reproductivamente maduros -ver figuras 2 y 3-. La venta es tanto de la semilla como del producto adulto procesado en pequeñas plantas. La mantención considera en hitos específicos

del ciclo de crecimiento del ostión, una actividad que se llama desdoblamiento, que consiste en la disminución de la densidad del producto en las jaulas para permitir el crecimiento apropiado de los individuos más aptos, además de la limpieza de otros organismos vivos que se adhieren. Los clientes son principalmente pequeñas empresas procesadoras que generan un producto final, las cuales igualmente se encuentran asociadas gremialmente. Adicionalmente hay pequeños comerciantes, que le dan muy poco valor agregado al producto, a lo más un incipiente envasado. Los proveedores de este negocio son principalmente: insumos de cultivos, linternas, bollas, pinturas para botes, lanchas, motores, mallas, jaulas, etc. Mucho de ellos son Mipymes.

En cuanto a la Salmonicultura Chilena, ésta es una industria muy exitosa que se ha transformado en tan solo veinte años en el segundo exportador mundial de salmón, manteniendo permanentemente tasas de crecimiento por sobre los dos dígitos. Como se establece en estudios anteriores [4], esta industria es muy competitiva a nivel mundial. En efecto, localmente se ha producido un encadenamiento hacia atrás con el surgimiento de muchas Mipymes que proveen de insumos, las cuales requieren innovar en tecnologías por la misma presión que ejerce la competitiva industria del salmón. Los productos y servicios involucrados en el modelo de negocios de la industria salmonera, incluyen la venta del salmón - atlántico, trucha y coho - en sus distintas etapas de cultivo. El ciclo constituido por estas etapas de cultivo, se encuentra directamente relacionado con el crecimiento de la biomasa, e intenta emular el ciclo natural del salmón. Contempla principalmente: etapa de piscicultura, etapa en lago, etapa en estuario, etapa en engorda.

Además, hay que considerar actividades previas asociadas al desarrollo genético de ovas. Por otro lado, hay actividades del proceso relativas al procesamiento en planta - clasificación, fileteo, embasado y estibamiento - y a venta, marketing y distribución. Por ejemplo, una pequeña proporción de los individuos se venden en la etapa de piscicultura, lo que se denomina venta small, contemplando principalmente la venta de ovas. Los niveles de producción de esta industria son tales que todos los días se verifican cosechas.

Como se indicó, un potencial para la incorporación de Software Libre se encuentra en los proveedores de la industria, pues en su mayoría corresponden a Mipymes, muchas veces con rezago tecnológico. Dentro de estos proveedores se encuentran: proveedores de alimentos para el salmón, proveedores de bins para el transporte de productos, fábricas de mallas, empresas mantenadoras y limpiadoras de mallas, transporte de especies vivas, transporte marítimo, transporte terrestre, empresas proveedores de servicios informáticos, etc. Habría que agregar que además se tienen los centros de cultivo, que corresponden a concesiones en el agua. Estos proveedores tienen la necesidad de apoyar sus procesos logísticos con software especializado principalmente en áreas tales como administración de la relación con el cliente, gestión de

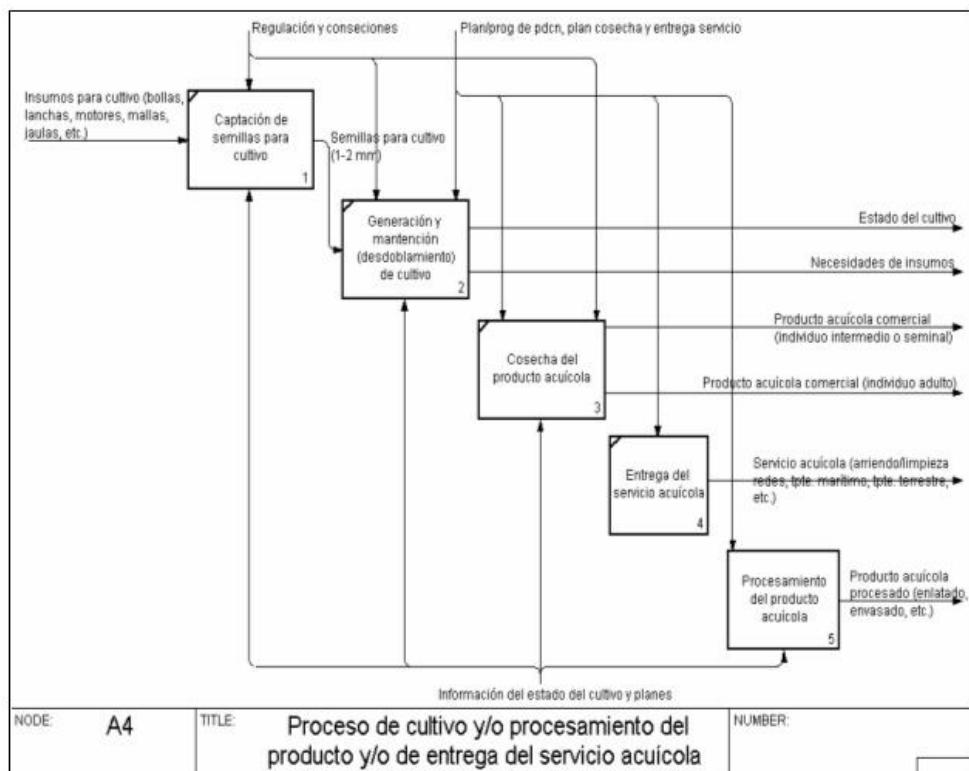


Figura 3. Proceso de Cultivo y/o Procesamiento del Producto y/o Entrega del Servicio Acuícola.

la producción y/o entrega del servicio y gestión de calidad. Con respecto a las empresas proveedoras de servicio informáticos, se conoció la operación en terreno de una de ellas, y se pudo constatar que existe un rezago considerable en metodologías y suites de apoyo al proceso de desarrollo de software.

Dado que muchas de las redes de la industria son inalámbricas por encontrarse los cultivos en el mar, es necesario proveer de software para el monitoreo de redes, entre otras necesidades específicas soportadas ampliamente por el Software Libre. Otros actores de la industria, de los cuales también se obtuvieron encuestas, le dan valor agregado al producto a través de envasado y enlatado, el que luego distribuyen principalmente a mercado internacional - Inglaterra, Tailandia, Malasia, Australia, etc. -, a través de relaciones con recibidores extranjeros. Estas empresas son Mipymes, con pequeños centros productivos, frigoríficos y cámaras de frío. Los abastecedores del producto congelado para su elaboración, son las grandes empresas salmoneras. El tipo de proveedores que tienen estos actores corresponden a: transporte terrestre, transporte marítimo, manutención de maquinaria de clasificación y enlatado, grúas horquilla, etc.

Las actividades genéricas del proceso productivo son: línea de limpieza del producto, moldeo y pesaje (incorporación del producto en las latas), tapado,

codificado, empaque y etiquetado. La programación de la producción es por pedido, y se encuentra primariamente apoyada por herramientas básicas como Excel. Se pudo constatar que no existe mayormente planificación a nivel táctico de la producción, por lo menos a base de estimaciones cualitativas de la demanda, y menos a nivel estratégico.

Las aplicaciones principales que se deben considerar para este sector, principalmente para los proveedores y pequeños productores, dicen relación con gestión de inventario, página Web de productos, gestión de clientes, gestión de proveedores y gestión de la distribución.

4.3. Sector Maderero

Los rubros esenciales de la industria maderera que entregan valor agregado al producto, son: aserrío, elaboración de maderas estructurales y revestimientos, vivienda y muebles de madera. Estas empresas son todas Mipymes que se abastecen del centro de acopio con maderas aserradas de la octava, novena y décima regiones. Su inventario es reducido, de maderas como Encina, Raulí, Pino y Álamo, además de Trupán y Cholguán. Las maderas finas - Roble, Cedro, Tepa, Lingue, Mañío, Coigüe, etc. - se abastecen preferentemente a pedido, y no se verifica stock para ellas.

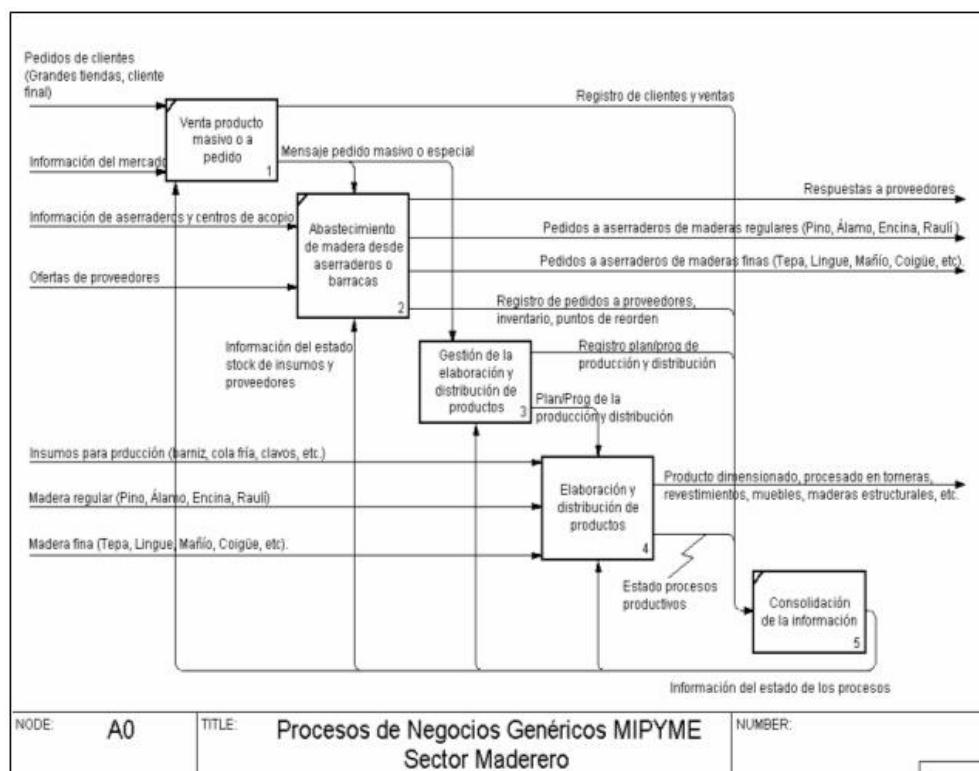


Figura 4. Procesos genéricos claves Mipymes Sector Maderero.

Los principales tipos de clientes son finales y grandes tiendas que trabajan bajo la estrategia de marcas propias. Por lo tanto, el proceso productivo que soporta la generación de productos para estos clientes es en masa. Sus procesos productivos utilizan maquinarias especializadas para corte, hornos de secado, barnizado, etc. Se verifica que un grupo de estas Mipymes trabajan en forma manual, es decir, son intensivas en mano de obra. Sin embargo, todas utilizan juicio experto para la programación de la producción.

Por otro lado se detecta un gran rezago en el apoyo tecnológico de sus procesos de negocios claves, tales como marketing y análisis de mercado; venta y atención de clientes; gestión de abastecimiento e inventarios y gestión de la entrega. En síntesis las aplicaciones principales que se deben considerar para este sector, dicen relación con gestión de inventario, página Web de productos y gestión de la producción.

5. Determinación de Requerimientos Tecnológicos Funcionales para las Mipymes.

Un conjunto de las actividades de los procesos de negocios de los sectores analizados pueden automatizarse y complementarse con buenas prácticas de gestión. En particular, los subprocesos de las actividades de los niveles superiores son los que definen los requerimientos tecnológicos funcionales. Las actividades que se muestran en el listado del anexo 1, constituyen el conjunto de subprocesos seleccionados que pueden apoyarse con herramientas tecnológicas adecuadas, y por lo tanto, definen los requerimientos funcionales de las Mipymes en los sectores analizados. Los requerimientos funcionales emanados de los subprocesos seleccionados pueden consultarse en la tabla del anexo 2. Un listado completo de actividades y requerimientos funcionales pueden consultarse en [21].

6. Selección de Herramientas Basadas en Software Libre.

Una vez determinados sistemáticamente en las secciones anteriores el rezago tecnológico de las Mipymes, sus procesos de negocios y los requerimientos tecnológicos derivados de éstos últimos, es hora de identificar las aplicaciones de software libre existentes para satisfacer tales requerimientos. Para ello, se utiliza la siguiente metodología de evaluación.

6.1. Metodología de Evaluación

Se realizó una categorización de las soluciones disponibles en dos grandes grupos funcionales: soluciones de infraestructura tecnológica y soluciones de negocios. Respecto a las soluciones de infraestructura tecnológica, se realizó una revisión y análisis de las aplicaciones de esta categoría disponibles bajo software libre en función de las necesidades y posibilidades reales de implementación en las Mipymes de los sectores en estudio. De acuerdo a esa información, las principales herramientas, dentro de esta categoría de soluciones son: (1) sistemas operativos (clientes y servidores), (2) correo electrónico, (3) sistema administrador de bases de datos, (4) servidor de aplicaciones (software), y (5) herramientas ofimáticas.

Respecto a las soluciones de negocios disponibles bajo software libre, se analizó y evaluó en función de las necesidades y procesos identificados en la sección anterior. Como parte de este análisis, se determinó si estas soluciones requieren algún nivel de adaptación y/o ajustes. Adicionalmente, se incorporaron una serie de criterios relacionados con la factibilidad de implementación y viabilidad en el tiempo de adopción de estas herramientas. Estos criterios son: (1) orientación a procesos (workflow), (2) esquema de operación (Web, cliente/servidor, etc.), (3) compatibilidad (sistema operativo, bases de datos, etc.), (4) utilización de estándares abiertos, (5) soporte comercial, (6) robustez del proyecto y de la comunidad, y (7) requerimientos técnicos y de capacitación. El detalle del análisis puede verse en [21].

6.2. Análisis de Aplicabilidad de Soluciones de Infraestructura Tecnológica.

Para este caso, existe bastante evidencia práctica de que se trata de soluciones que son factibles de implementar, en lo técnico [21], [6], [8], [9], [10], [11], [12], [13], [14]. Sin embargo, existen una serie de factores adicionales que determinan la factibilidad práctica de implementación de este tipo de soluciones en las Mipymes. Dentro de los factores detectados, los que se utilizan para el análisis son:

- *Grado de complejidad en la capacitación de usuarios y técnicos en las soluciones* (G1): se utiliza una escala de 1 a 5, donde 1 representa mayor nivel de complejidad y 5 un menor nivel de complejidad.
- *Grado de disponibilidad de soporte técnico requerido para la implementación de la infraestructura* (G2): se utiliza una escala de 1 a 5, donde 1 representa menor disponibilidad y 5 representa una mayor disponibilidad de soporte técnico.

- *Grado de complejidad en la migración (de datos, configuraciones, etc.) o configuración respecto a soluciones propietarias de uso general (G3):* se utiliza una escala de 1 a 5, donde 1 representa mayor nivel de complejidad y 5 un menor nivel de complejidad.

Con estos valores se determina un puntaje, calculado como la multiplicación de los tres grados indicados anteriormente, $G_1 \times G_2 \times G_3$. Respecto a la factibilidad de implementación de estas soluciones, se utiliza la siguiente categorización, en base al puntaje obtenido por cada herramienta:

- *Nivel 0:* no es factible la implementación de este tipo de soluciones en las Mipymes de los sectores en estudio. Corresponde a soluciones con puntajes menores a 10.
- *Nivel 1:* existe un nivel bajo de factibilidad de implementación. Corresponde a soluciones mayores de 10 y menores de 25 puntos.
- *Nivel 2:* existe un nivel medio de factibilidad de implementación. Corresponde a soluciones mayores o iguales de 25 y menores de 60 puntos.
- *Nivel 3:* existe un nivel alto de factibilidad de implementación. Corresponde a soluciones mayores o iguales de 60 puntos.

A continuación se presenta el resultado de dicho análisis⁵:

Necesidad	Herramienta tecnológica	Soluciones de software libre disponibles	Grado de factibilidad de implementación
1 Estación de Trabajo	Sistema Operativo	Ubuntu Linux	Nivel 2
2 Estación de Trabajo	Navegador de Internet	Mozilla Firefox	Nivel 3
3 Estación de Trabajo	Herramientas de Ofimática	OpenOffice	Nivel 3
4 Estación de Trabajo	Cliente de Correo Electrónico	Mozilla Thunderbird	Nivel 2
5 Estación de Trabajo	Cliente de Mensajería Instantánea	Gaim	Nivel 2
6 Servicios de Red	Servidor de Correo Electrónico	Postfix / Sendmail	Nivel 1
7 Servicios de Red	Filtro Anti - spam	DSPAM / SpamAssassin	Nivel 0
8 Servicios de Red	Cortafuegos	Shorewall / Smoothwall	Nivel 0
9 Servicios de Red	Solución de Directorios y Archivos e Impresoras Compartidos en Red	Samba	Nivel 1

Tabla 3. Aplicación de Soluciones de Infraestructura Tecnológica.

6.3. Análisis de Aplicabilidad de Soluciones de Negocios.

Al igual que en el caso de las soluciones de infraestructura, se consideran los tres factores y categorizaciones del puntaje. A continuación se presenta el resultado de dicho análisis, por sector y segmento de empresa.

⁵Detalle exhaustivo en [21].

6.3.1. Microempresa

En este segmento de empresas, la principal complejidad es el bajo nivel de recursos disponibles para la implementación de un proyecto tecnológico, puesto que se trata de instituciones que en muchos casos cuentan con apenas 2 a 4 empleados, casi todos socios. Sin embargo, el trabajo en conjunto es una interesante alternativa a considerar, puesto que es posible agrupar recursos de forma de abordar colaborativamente proyectos de mayor envergadura y presupuesto. Además estos proyectos desarrollados en forma asociativa son la forma usual como se desarrollan los proyectos de software libre.

A continuación se muestra como ejemplo el resultado obtenido para el sector agrícola/frutícola -ver tabla 4-. Los dos sectores restantes pueden consultarse en [21].

Necesidad	Herramienta tecnológica	Soluciones de software libre disponibles	Grado de factibilidad de implementación
Apoyo a Procesos Críticos	Herramientas de Ofimática	OpenOffice (Writer, Calc y Base)	Nivel 3
Apoyo a Implementación Mejores Prácticas Norma Nch2909	Herramientas de Ofimática	OpenOffice (Writer, Calc y Base)	Nivel 3
Construcción de Página Web de Información de la Empresa	Herramientas de Ofimática	OpenOffice Writer/Web	Nivel 3

Tabla 4. Aplicación de Soluciones de Negocio a Microempresas del Sector Agrícola / Frutícola.

6.3.2. Pequeña Empresa

En este segmento se encuentran empresas de entre 5 y 50 empleados, incluido a los socios, con posibilidad e interés por invertir en soluciones tecnológicas de bajo costo, pero de impacto apreciable. Al igual que en el caso anterior, se considera que si bien una solución tecnológica podría ser costosa para una empresa, no lo sería para una comunidad de éstas, las cuales estarían dispuestas a considerar el financiamiento parcial de la implantación de una tecnología que beneficie a todo el conjunto (caso pequeñas empresas de producción de ostiones en Tongoy, las cuales manifestaron voluntad de trabajar en conjunto la implantación de una solución tecnológica)

	Necesidad	Herramienta tecnológica	Soluciones de Software Libre Disponibles	Grado de Factibilidad de Implementación
1	Construcción de Sitio Web Corporativo	Administrador de Contenidos Web	Joomla	Nivel 3
2	Construcción de Sitio Web de Catálogo de Productos	Plataforma de Comercio Electrónico	OsCommerce	Nivel 2
3	Sistema de Gestión de Inventario	Herramienta de Automatización de Procesos (Workflow)	Apache Ofbiz	Nivel 1

Tabla 5. Aplicación de Soluciones de Negocio a Pequeñas Empresas del Sector Agrícola / Frutícola.

6.3.3. Mediana Empresa

Respecto al grado de adaptación requerido de las tecnologías de software, se utilizó la siguiente categorización ⁶:

- *Nivel 0*: no existe una solución de software libre para el proceso.
- *Nivel 1*: existe una solución de software libre, pero requiere de mucha adaptación para el caso chileno (70 % de adaptación o mas).
- *Nivel 2*: existe una solución de software libre, pero requiere de un 50 % de adaptación.
- *Nivel 3*: existe una solución de software libre y solo requiere de un 30 % de adaptación.
- *Nivel 4*: existe una solución de software libre y casi no hay que adaptarla, solo se requieren ajustes funcionales menores y configuración.

7. Conclusiones

En este paper se muestra el estado actual del uso de TICs por parte de las Mipymes nacionales de los sectores agrícola/frutícola, acuícola y maderero. Además se describe su disposición a implantar software libre, los procesos de negocios claves que requieren apoyo tecnológico, los requerimientos funcionales emanados de éstos y la selección de soluciones tecnológicas basadas en software libre que los satisfacen. Se pudo constatar que los empresarios de las Mipymes,

⁶El resultado de la selección de soluciones de negocio a medianas empresas basadas en software libre, para el caso agrícola/frutícola puede consultarse en el anexo 2.

en general, dan una importancia estratégica al uso de las TICs en sus respectivos negocios. Adicionalmente, se probó lo que normal y cualitativamente se observa como el rezago tecnológico de las Mipymes Chilenas. Sin embargo, no realizan un mayor nivel de inversión en estas materias por desconocimiento, falta de recursos o simplemente porque no han encontrado aplicaciones tecnológicas que se adapten a sus necesidades particulares, manifestando el temor de que el proceso se haga al revés, es decir, que el negocio se deba adaptar a la solución.

Las empresas aprueban la idea de conformar un proyecto de implantación de estas aplicaciones. Un caso relevante en este sentido es el gremio del cultivo del ostión en el sector acuícola, comprometiendo al menos horas hombre como contraparte usuario de las herramientas. Esto permitiría en el futuro dar curso a un proyecto piloto, el cual ya ha sido diseñado y se puede consultar en [21].

De las soluciones de software libre existentes, en este paper se distinguen dos grandes categorías: infraestructura tecnológica y de negocios. La primera se refiere a aquellas soluciones que implementan la base tecnológica deseada para la correcta ejecución de la Mipyme y que posibilita la implantación de otras soluciones más complejas. La segunda se trata de aplicaciones de software libre directamente relacionadas con la automatización de un proceso de negocio de la Mipyme.

En el ámbito de las soluciones de infraestructura tecnológica, existe un conjunto de herramientas de software libre que pueden utilizarse prácticamente sin ninguna modificación o adaptación, y que requieren principalmente de capacitación de los usuarios y soporte para ser factibles de implantar.

En el caso de las soluciones de negocio, a partir del modelamiento de los procesos críticos detectados en las Mipymes, se concluye que NO existen soluciones de negocio disponibles como software libre que se puedan aplicar directamente o cumplan al 100 % las necesidades de estas empresas. Sin embargo, algunas de estas herramientas pueden ser utilizadas como base, adaptadas y transformarse en una solución concreta a las necesidades y entorno de las Mipymes nacionales.

Respecto a la factibilidad del uso de herramientas de software libre en las Mipymes, se concluye que al menos en lo que a infraestructura tecnológica se refiere, su implantación no debería presentar mayor problema. El verdadero desafío se concentra en las soluciones de negocios, siendo altamente recomendable el desarrollo de un proyecto piloto para la implantación de estas herramientas en el futuro.

Si bien es cierto que el software libre permite ahorros significativos respecto a adquisición de licencias, su implementación está sujeta a las restricciones de costos de cualquier proyecto tecnológico. En efecto, el recurso humano utilizado es similarmente costoso al caso del software comercial, por lo cual el proyecto de implantación debe considerar los costos asociados a adaptación, capacitación,

implementación, soporte, etc.

La experiencia internacional muestra que los proyectos de implantación de software libre, en el ámbito de las soluciones de negocios, presentan una mayor probabilidad de éxito en la medida que se constituye un grupo de trabajo que involucre a varias empresas del mismo rubro y con similares necesidades, configurándose así un cluster a partir del cual se generará una experiencia colectiva que permita extraer las mejores prácticas en el desarrollo de este tipo de proyectos y aprovechar economías de escala.

Actualmente, la investigación realizada por los autores ya ha dado su primer fruto por medio de cursos para microempresarios por parte de Sofofa, que se han estado dictando desde Julio de 2007, quienes se basaron en las principales necesidades y aplicaciones determinadas en esta investigación [19]. Además, se encuentra en desarrollo un proyecto de implantación en el sector acuícola.

Agradecimientos: Esta investigación se efectuó en el marco del estudio encargado por la Japan Official Development Assistance, el Programa de las Naciones Unidas para el Desarrollo (PNUD), y la Sociedad de Fomento Fabril (SOFOFA). Los autores además agradecen a Diana Fuenzalida Tapia y a Magdalena Carrasco, por su participación en la elaboración de encuestas telefónicas a algunas Mipymes.

Referencias

- [1] Aral, Sinan; Brynjolfsson, Erik and Wu, D.J. "Which Came First, IT or Productivity? The Virtuous Cycle of Investment and Use in Enterprise Systems". MIT Center for Digital Business Working Paper. October, 2006.
- [2] Brynjolfsson, Eric; Lorim M. Hitt. "Beyond the Productivity Paradox, Computers are the catalyst for bigger changes". MIT Sloan Management, 1998.
- [3] Carr, Nicholas G. "La TI ya no es importante". Harvard Business Review, 2004.
- [4] Eduardo Bitrán C. "Crecimiento e Innovación en Chile". Revista Iberoamericana de Ciencia, Tecnología, Sociedad e Innovación. Organización de Estados Iberoamericanos Para la Educación, la Ciencia y la Cultura. Número 4. Septiembre - Diciembre 2002.
- [5] Iván Coydan T, Álvaro Rojas M y José Díaz O. "El Cambio Estructural en la Agricultura Chilena: Tenencia de la Tierra y Empleo Agrícola". Universidad de Talca - CHILE.

- [6] <<http://www.chileforge.cl>>
- [7] Dorgan S.; Dowdy J. "When IT Lifts Productivity: Companies Should Beef up Their Management Practices Before Focusing on Technology". The McKinsey Quarterly: The Online Journal of McKinsey & Co. 2004 Number 4.
- [8] <<http://www.tux.cl>>
- [9] <<http://www.sourceforge.net>>
- [10] <<http://chileforge.cl/projects/pygestor/>>
- [11] <<http://www.rizoma.cl>>
- [12] <<http://www.phpmyfactura.cl>>
- [13] <<http://remunex.chileforge.cl/home-page/>>
- [14] <<http://pymegnu.sourceforge.net>>
- [15] <<http://www.sofofa.cl>>
- [16] <<http://www.chilealimentos.cl>>
- [17] <<http://www.lignum.cl>>
- [18] <<http://www.dapel.cl>>
- [19] <<http://www.sofofa.cl/sofofa/index.aspx?channel=4319&appintanceid=12189&pubi>>
- [20] I. Witten, E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Junio 2005
- [21] <<http://200.27.194.5/sofofa/index.aspx?channel=4319>>. También consultar: <<http://wi.dii.uchile.cl/noticias/informe-uso-software-libre-en-mipymes>>
- [22] <www.idef.com>

8. ANEXOS

Anexo 1: Listado de Procesos y Subprocesos.

		Proceso	ID Subproceso	Subproceso
ID Proceso	FP ₁	Ventas y comercialización (mercado externo y nacional)	FP ₁ S ₁	Subproceso gestión de clientes y pedidos externos.
	FP ₂	Administración de abastecimiento de productos agrícolas/frutícolas	FP ₁ S ₂	Subproceso gestión de clientes y pedidos mercado nacional.
			FP ₂ S ₁	Subproceso registro de estimaciones de cosecha.
			FP ₂ S ₂	Subproceso planificación de entregas de bins.
	FP ₃	Administración de insumos	FP ₂ S ₃	Subproceso obtención de reportes de cosecha y entregas.
			FP ₃ S ₁	Subproceso generación de pedidos.
	FP ₄	Gestión del procesamiento de productos agrícolas/frutícolas	FP ₃ S ₂	Subproceso gestión del inventario.
			FP ₄ S ₁	Subproceso planificación agregada de las operaciones.
	FP ₅	Procesamiento de productos agrícolas/frutícolas	FP ₄ S ₂	Subproceso programación diaria y por turnos de las operaciones.
			FP ₅ S ₁	Subproceso registro del estado del proceso.
			FP ₆ S ₁	Subproceso operaciones unitarias iniciales
			FP ₆ S ₂	Subproceso congelado o deshidratado
	FP ₆	Elaboración de frutas (congelados deshidratados)	FP ₆ S ₃	Subproceso operaciones unitarias finales
			FP ₆ S ₄	Subproceso control de calidad

Tabla 6. Procesos/Subprocesos Sector Frutícola / Agrícola.

Anexo 2: Aplicación de Soluciones de Negocio a Medianas Empresas

	Proceso/ Subproceso	Requerimiento	Soluciones de Software Libre Disponibles	Grado de Adaptación de Tecnologías de Software
1	FP1S1	CRM con módulo comercial para exportaciones.	OpenCRX	Nivel 2
2	FP1S2	CRM con módulo comercial para mercado interno.	OpenCRX	Nivel 3
3	FP2S1	Módulo de abastecimiento que permita registrar estimaciones de producción de productos agrícolas/frutícolas, ordenados por especie y variedad.	-	Nivel 0
4	FP2S2	Módulo de distribución que permita asignar a huertos específicos la entrega en una fecha determinada de los bins vacíos, ruteando óptimamente con heurísticas simples (o saber experto) la ruta entre distintos huertos. Lo mismo se requiere para el proceso de "recogida" de bins llenos con producto cosechado.	-	Nivel 0
5	FP2S3	Módulo para generar reportes de gestión filtrados por fechas, producto, variedad, etc. de las cosechas y entregas de producto.	JasperReports	Nivel 3
6	FP3S1	Sistema de gestión de inventarios (órdenes de compra a proveedores, rebaja de stock, lote económico de pedido)	Apache Ofbiz	Nivel 3
7	FP3S2	Sistema de gestión de inventarios (órdenes de compra a proveedores, rebaja de stock, lote económico de pedido)	Apache Ofbiz	Nivel 3
8	FP4S1	Sistema para generar plan de producción (cruce de variables: producto, variedad, cosecha estimada, fecha de cosecha, fecha de procesamiento, volumen de procesamiento)	-	Nivel 0
9	FP4S2	Sistema para generar programa diario o por turnos del procesamiento (orden óptimo de ingreso de los lotes a las máquinas clasificadoras)	-	Nivel 0
10	FP5S1	Sistema para generar programa diario o por turnos del procesamiento (debe permitir registro del estado del proceso en puntos específicos del proceso)	-	Nivel 0

MÉTODOS CUANTITATIVOS PARA UNA NUEVA CONFIGURACIÓN DE UNIDADES TERRITORIALES EN LICITACIONES DE JUNAEB

GONZALO ZAMORANO*
GUILLERMO DURÁN*
ANDRÉS WEINTRAUB*

Resumen

La Junta Nacional de Auxilio Escolar y Becas (JUNAEB) es un organismo gubernamental de Chile que tiene como misión ayudar a niños y jóvenes con alta vulnerabilidad social a su inserción y permanencia en el mundo escolar. Dentro del área alimenticia decenas de empresas se encargan de suministrar las comidas a los colegios del país, el cual estaba dividido hasta 2007 en 136 Unidades Territoriales (UT). La composición de estas unidades había sido elaborada hace varios años de manera manual principalmente por restricciones geográficas y, en menor medida, tomando en cuenta la cantidad de raciones alimenticias, lo que generaba UT con atractivos muy diferentes para las empresas que año a año se presentan a las licitaciones que organiza JUNAEB. Las conversaciones con autoridades de la institución y directivos de las empresas nos permitieron concluir que existen varios factores que influyen en el atractivo de un territorio, como por ejemplo, el mencionado número de raciones, el número de colegios, la superficie de la UT y la accesibilidad a los colegios. El trabajo desarrolla diferentes métodos de programación matemática para determinar una nueva configuración a ser empleada en el proceso de licitación, de modo que las UT de una misma región sean similarmente atractivas para las empresas, y que gracias a ello todos los colegios de una misma región tengan la posibilidad de una alimentación de buena y similar calidad. Esto permite agregar mejoras al actual proceso, tanto operativas como económicas. Los supuestos considerados para la elaboración de la nueva configuración incluyen políticas de JUNAEB y algunas peticiones empresariales. También se busca facilitar el cálculo de los costos de cada oferta, relacionándolos con los criterios usados. A partir de la licitación de 2007 JUNAEB está utilizando una de las configuraciones de unidades territoriales propuestas en este trabajo.

*Departamento de Ingeniería Industrial, Universidad de Chile

1. Introducción

La Junta Nacional de Auxilio Escolar y Becas (JUNAEB) es un organismo del gobierno de Chile cuyo fin es ayudar a los niños y jóvenes con alta vulnerabilidad social a su inserción y mantención en el sistema educativo. Para ello desarrolla un conjunto de programas en numerosas áreas, tales como alimentación, salud escolar, salud mental, salud oral, vivienda, recreación, útiles escolares y becas, orientados a entregar una asistencia integral que haga efectiva la igualdad de oportunidades ante la educación. Su acción llega a casi dos millones de estudiantes de la educación pública en forma diaria, en establecimientos de todo el país.

El programa de alimentación escolar (PAE) es brindado a cerca de 10 mil establecimientos municipalizados y particulares subvencionados a lo largo de todo Chile. Consiste en el suministro a las Unidades Territoriales (UT) en las cuales es dividido el país, de distintos tipos de comida para los niños y jóvenes en edad escolar. La distribución de alimentos a estas unidades es realizada por compañías dedicadas al rubro alimenticio.

En Chile hay 346 municipalidades, de las cuales 104 (30 %) atienden a comunas de menos de 10.000 habitantes (44 de ellas son comunas de menos de 5.000 habitantes); 195 (57 %) son comunas de entre 10.000 y 100.000 habitantes y 46 (13 %) son comunas de más de 100.000 habitantes. JUNAEB asocia estas comunas en UT a lo largo de todo Chile. Hasta 2007 había 136 de estas unidades, sin embargo 34 de ellas no fueron analizadas en este estudio, puesto que JUNAEB las tiene destinadas a empresas Pymes y no pretende modificarlas. Las Figuras 1 y 2 muestran la distribución de las UT en el país antes de los cambios hechos en este estudio. Para el desarrollo de este trabajo se determinó que los valores mínimo y máximo del número de raciones de cada UT sería de 15.000 y 40.000, respectivamente. Estos valores fueron consensuados con JUNAEB y obedecen, en el caso de la cota inferior, a cuestiones operativas de JUNAEB y a la dimensión mínima requerida para que las firmas obtengan apropiadas economías de escala, y en el caso de la cota superior, al deseo de la institución de evitar que haya regiones con presencia de muy pocas empresas.

La información necesaria para cuantificar cada criterio de los considerados en este trabajo corresponde al mes de abril de 2007. Si bien estos datos pueden cambiar durante el transcurso de los meses, el resultado no debiera variar mayormente con estos cambios, pues en el corto plazo las comunas mantienen (o varían levemente) su importancia relativa al compararse con otra de la misma región. Por ejemplo, si las raciones a nivel país aumentan en un 5 %, también aumentarán en una cifra similar la cantidad de raciones en cada comuna. Lo

mismo ocurre con el resto de los criterios a ser considerados.

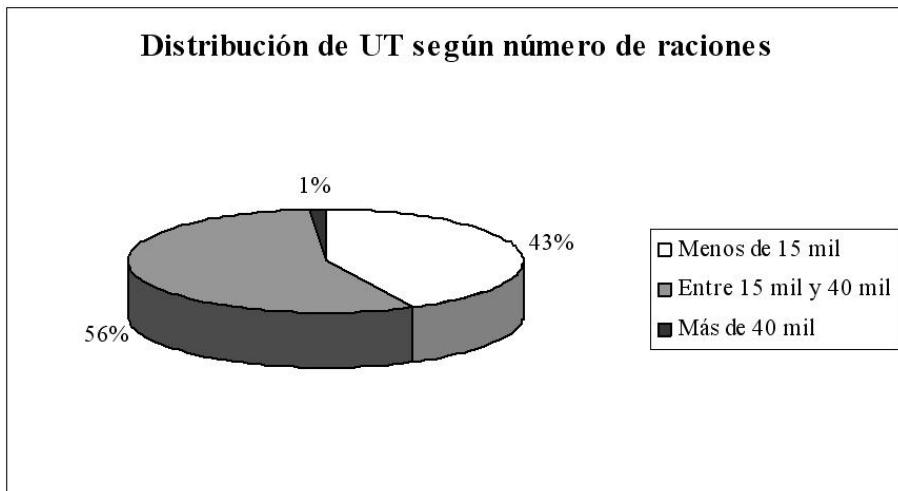


Figura 1. Distribución de UT hasta 2007 según número de raciones.

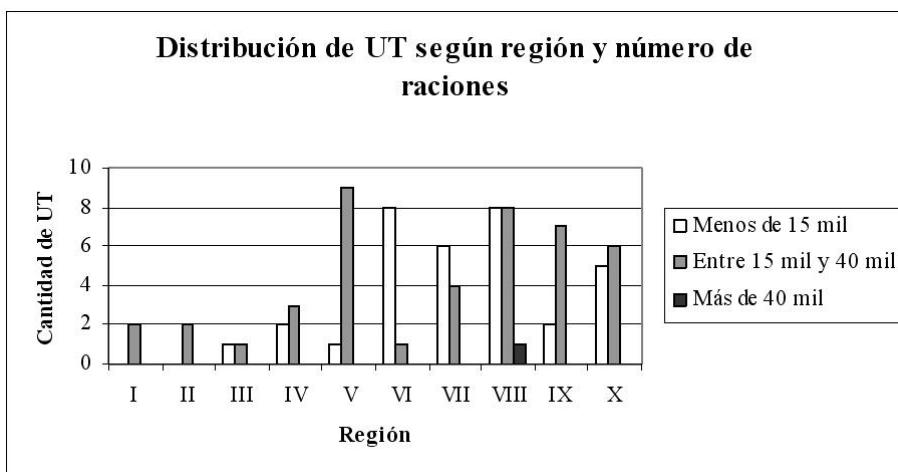


Figura 2. Distribución de UT hasta 2007 según región y número de raciones.

Como se observa en la Tabla 1, la tendencia indica que mientras mayor es el número de raciones a suministrar, mayor es el número de territorios en el que es dividida la región. Este hecho permite concluir que la cantidad de raciones era un aspecto importante en la distribución de las comunas para la conformación de las UT.

El gran problema que tenía la configuración territorial utilizada hasta 2007 para las licitaciones, tanto para las empresas como para JUNAEB, era la disparidad de atractivos entre las UT, ya que sólo era tomado en cuenta un criterio geográfico y de cierta equidad en el número de raciones en cada una de las unidades de una región. En la elaboración de los territorios se había tratado

Región	Nº UT	Total Raciones/Día
I	2	50,350
II	2	39,844
III	2	35,114
IV	5	87,777
V	10	186,941
VI	9	104,187
VII	10	157,525
VIII	17	286,860
IX	9	158,915
X	11	197,087
XI	1	10,541
XII	1	12,177
RM	57	655,429
Total	136	1,982,747

Tabla 1. Cantidad de raciones y UT por región hasta 2007. Fuente JUNAEB, abril 2007.

de aislar a las comunas con un gran número de raciones, asignándolas solitariamente a una UT. Para el resto de la región se había buscado equiparar esta cantidad de alimentos, no importando la superficie que abarcara esa unidad o el número de colegios que se necesitara para alcanzar tal cifra.

El fin central del problema abordado en [7] y en el presente trabajo es obtener la estructura territorial de comunas más adecuada para conseguir unidades relativamente similares, es decir, disminuir la brecha entre los territorios atractivos y los no atractivos para las empresas. De esta manera se puede lograr que todos los colegios de una misma región tengan la posibilidad de una alimentación de buena y similar calidad, evitando además que las empresas amarren en sus ofertas unidades atractivas con otras que no lo son, distorsionando el valor de las ofertas (dado que éstas se adjudican completas para aprovechar las economías de escala).

La generación de unidades o bloques para la resolución de problemas se da también en otros ámbitos. Podemos citar el caso forestal, cuando se quiere desarrollar una planificación de cosechas, respetando un máximo de área contigua que puede ser cosechada en un período (ver por ejemplo [2,3,4]); o el caso de asignación a escuelas de estudiantes que viven en sectores de una ciudad, respetando ciertos criterios (ver por ejemplo [1,6]).

En nuestro trabajo, el atractivo de una UT será obtenido gracias a un puntaje compuesto por los aspectos más relevantes que tiene una comuna, tanto para JUNAEB como para las empresas concesionarias: raciones, colegios, superficie y accesibilidad. El número de raciones da una idea cabal de la magnitud del negocio; el número de colegios es importante por el costo fijo de atender a cada uno; la superficie que se recorre influye por los costos de traslados y transporte; mientras que la presencia de colegios de difícil acceso

en una UT también aumenta los costos de las empresas.

Cada comuna posee una serie de datos que la identifican, los cuales brindan información sobre características de sus colegios, número de raciones, tipos de raciones, a qué UT pertenecen, etc. De esta forma, generalmente la unidad básica del problema es la comuna, pues está completamente caracterizada, aunque puede darse en algún caso que la misma se divida en dos o más partes.

Al tener este problema cuatro objetivos que pueden ser conflictivos entre sí, surge la necesidad de aplicar técnicas de objetivos múltiples. Para este trabajo se eligió una variación del Proceso de Jerarquía Analítica (AHP), creado por Thomas Saaty [5], para determinar el peso que tiene cada criterio. Este índice, unido al peso que tiene una UT por criterio en cada región, permite asignar un puntaje a los posibles territorios a ser seleccionados.

Para encontrar soluciones de este problema se desarrollan tres metodologías. En primer lugar es utilizado un algoritmo heurístico de búsqueda local para intercambiar comunas entre UT, con el objetivo de minimizar la desviación estándar de los puntajes y mejorar la configuración inicial. Como segunda metodología se desarrolla un algoritmo de elaboración de clusters de comunas para configurar cada UT (usando ideas de [3]), sujeto a ciertas restricciones. A cada cluster se le calcula luego un puntaje. Clusters y puntajes son usados como inputs dentro de un modelo exacto de programación lineal entera que minimiza la diferencia entre el mayor y menor puntaje en cada región. Este objetivo arroja resultados concordantes con el objetivo de minimizar la desviación estándar en la mayoría de los casos. Como última metodología también se busca minimizar la desviación estándar utilizando el algoritmo heurístico, pero tomando como solución inicial el resultado entregado por el modelo de programación entera.

Este estudio entrega una solución para 10 de las regiones de Chile existentes hasta 2006, buscando cambiar y mejorar la configuración inicial del país. En general, las soluciones entre las tres metodologías empleadas no distan mucho entre sí pero mejoran sustancialmente la situación existente hasta 2007. La primera parte del estudio que se finalizó fue la nueva configuración propuesta por la heurística, y por ello a partir de 2007 JUNAEB está utilizando esta solución en su proceso de licitación que abarca tres años (un tercio del país en cada año). La finalización a posteriori del mismo estudio para las otras dos metodologías permite afirmar que la solución empleada por JUNAEB es de muy buena calidad.

2. Importancia Relativa de una UT

La idea del procedimiento que se describe a continuación es encontrar un indicador que permita cuantificar el atractivo que tiene para las empresas un territorio dentro de una región; de esta manera se podrá analizar cuánto más atractiva es una UT que otra. En todos los casos, j representa a una UT y r a una región. La función que permite determinar el grado de atractivo de una UT tiene la siguiente forma:

$$U_{j,r} = a_{Rac}x_{Rac,j,r} + b_{Col}x_{Col,j,r} + c_{Sup}x_{Sup,j,r} + d_{Acc}x_{Acc,j,r}$$

$U_{j,r}$ =puntaje total que tiene la UT j de la región r.

a_{Rac} =importancia que tiene el número de raciones dentro de la canasta de criterios.

$x_{Rac,j,r}$ =peso que tiene la UT j de la región r según el número de raciones.

b_{Col} =importancia que tiene el número de colegios dentro de la canasta de criterios.

$x_{Col,j,r}$ =peso que tiene la UT j de la región r según el número de colegios.

c_{Sup} =importancia que tiene la superficie de la UT dentro de la canasta de criterios.

$x_{Sup,j,r}$ =peso que tiene la UT j de la región r según su superficie.

d_{Acc} =importancia que tiene el tipo de acceso dentro de la canasta de criterios.

$x_{Acc,j,r}$ =peso que tiene la UT j de la región r según el acceso a colegios.

La obtención de los valores $a_{Rac}, b_{Col}, c_{Sup}$ y d_{Acc} , usados en todas las metodologías, se hizo a través de una variación del Proceso de Jerarquía Analítica (AHP: Analytic Hierarchy Process) [5]. Este es un método que sirve para evaluar alternativas de decisión y elegir la mejor, cuando el tomador de decisión tiene múltiples objetivos o criterios, encontrando valores numéricos para cada alternativa. Por ejemplo, cuando una persona quiere comprar una casa puede tener varias alternativas posibles para elegir, mientras que puede tener diferentes criterios de selección como el costo (A), el número de habitaciones (B) o el barrio de la vivienda (C).

El AHP se puede resumir en los siguientes pasos:

1. Desarrollo de una comparación entre pares de alternativas de decisión por cada criterio, obteniendo una matriz de comparación por cada uno de éstos. Los niveles de preferencia para comparar las alternativas de decisión se muestran en la Tabla 2. En el ejemplo mencionado de la

compra de la casa, se compara las alternativas analizando el criterio A. Luego el B y el C, obteniendo tres matrices de comparación.

Nivel de preferencia	Valor
Igualmente preferido	1
Igualmente a moderadamente preferido	2
Moderadamente preferido	3
Moderada a fuertemente preferido	4
Fuertemente preferido	5
Fuerte a muy fuertemente preferido	6
Mu fuertemente preferido	7
Muy fuerte a extremadamente preferido	8
Extremadamente preferido	9

Tabla 2. Escala de Preferencias para comparación de a pares.

2. Proceso de Síntesis de las matrices, el cual consiste en los siguientes pasos:
 - Suma de los valores de cada columna de la matriz de comparación.
 - Normalización de las matrices, es decir, dividir cada elemento por la suma de la columna correspondiente.
 - Promedio de los valores de cada fila de las matrices normalizadas obteniendo los vectores de preferencia.
 - Combinación de los vectores de preferencia de cada criterio en una matriz de preferencia, la cual muestra la preferencia de cada alternativa para cada criterio.
3. Desarrollo de una comparación de a pares entre los criterios de decisión.
4. Proceso de Síntesis de la matriz de a pares de los criterios, el cual consiste en los siguientes puntos:
 - Suma de los valores de cada columna de la matriz de comparación de a pares de los criterios.
 - Normalización de la matriz, es decir, dividir cada elemento por la suma de la columna correspondiente.
 - Promedio de los valores de cada fila de la matriz normalizada. Éste sería el vector de preferencia de los criterios.
5. Multiplicación de la matriz obtenida en 2 por el vector obtenido en 4 para obtener los valores finales de cada alternativa.

La diferencia entre el AHP y lo realizado en el presente estudio es que sólo la comparación de los criterios es usada (Pasos 3 y 4 del resumen de AHP

recientemente definido), lo que permite cuantificar el atractivo de cada uno de ellos. La matriz de preferencias de los Pasos 1 y 2 es reemplazada por los puntajes según cada criterio que tiene cada UT (Secciones 3.1 o 4.2, según el método de resolución empleado).

Criterios	Raciones	Colegios	Superficie	Acceso	Peso
Raciones	1.00	1.00	3.00	3.00	38.07%
Colegios	1.00	1.00	2.00	3.00	34.22%
Superficie	0.33	0.50	1.00	2.00	16.94%
Acceso	0.33	0.33	0.50	1.00	10.77%
Suma	2.67	2.83	6.50	9.00	100.00%

Tabla 3. Relaciones entre los criterios y peso final de cada uno.

La Tabla 3 muestra la matriz que compara los criterios de a pares. Esta comparación fue realizada con el asesoramiento de funcionarios de JUNAEB. Por ejemplo, el valor 3 del elemento 3.1 de la matriz (fila 3, columna 1) refleja que el criterio raciones es “*moderadamente preferido*” sobre el criterio superficie (ver Tabla 2). Luego de realizar todas las comparaciones se procede a sumar los valores para cada criterio (Fila Suma). Para obtener el peso porcentual final por criterio, se promedian las importancias relativas de cada comparación, como se muestra en este ejemplo, donde se obtiene el peso porcentual de las raciones:

$$38,07 = \left(\frac{\frac{1,00}{2,67} + \frac{1,00}{2,83} \frac{3,00}{6,50} \frac{3,00}{9,00}}{4} \right) * 100$$

De este modo, los valores finales de cada criterio se encuentran indicados en la última columna de la Tabla 3. Las raciones tienen una importancia ponderada de un 38,07 %, confirmando que es el criterio más importante, lo que ya era considerado tanto por las empresas como por JUNAEB. Estos pesos fueron validados, al contrastar el valor de atractivo de las UT dado por estos criterios en la situación existente hasta 2007 con el análisis hecho por JUNAEB que determina cuán preferida es una UT. Es decir, los valores dados por estos criterios para el escenario original coinciden con los supuestos hechos por JUNAEB y las empresas acerca de UT atractivas y no atractivas. Con esta observación es posible concluir que los valores arrojados para cada criterio por el AHP son consistentes, y la comparación de a pares concuerda con lo que representa cada uno de los criterios para JUNAEB y las empresas. De esta forma es esperable que los resultados cuantitativos de las nuevas propuestas entreguen valores que expliquen fehacientemente la homogenización de cada región.

Los pesos de cada UT por criterio $x_{Rac,j,r}$, $x_{Col,j,r}$, $x_{Sup,j,r}$ y $x_{Acc,j,r}$ en cada región, analizados en secciones posteriores, son calculados de manera diferente

según la metodología ocupada (heurística o modelo de programación entera), aunque se mantiene una similar proporcionalidad al comparar UT en ambos casos.

Decimos que un criterio es directamente proporcional si mientras más proporción se tiene más atractiva es la UT; caso contrario ocurre con uno inversamente proporcional, en donde mientras más proporción se tenga menos atractiva es la UT. En el caso de los primeros se encuentra el número de raciones y accesibilidad, mientras que colegios y superficie son parte del segundo grupo.

3. Heurística de Búsqueda Local

El objetivo de esta heurística es encontrar un conjunto de UT de atractivo homogéneo en cada región. Como medida de homogeneidad se tomará la desviación estándar, es decir, nuestro algoritmo heurístico tratará de minimizar dicho valor (claramente, el valor ideal es una desviación estándar igual a 0).

La heurística utiliza para el intercambio de comunas un procedimiento de "búsqueda local", el cual comienza en una solución inicial (la configuración de cada región hasta 2007) y busca una dirección de máximo descenso del valor de la función objetivo (la desviación estándar), moviéndose a soluciones vecinas. Para ello se define un concepto de vecindad entre soluciones que se explicará más adelante.

3.1. Pesos de cada UT por criterio utilizados en la heurística

Analicemos la forma de calcular el peso de cada UT según cada criterio ($x_{Rac,j,r}$, $x_{Col,j,r}$, $x_{Sup,j,r}$ y $x_{Acc,j,r}$) para esta metodología. Notar que al momento de hacer los cálculos se conoce la partición en UT de toda la región. Para el ejemplo a considerar a continuación supondremos que la región está particionada en 2 UT: A y B.

■ Número de Raciones

El cálculo del peso en cualquier UT según este criterio consistió en una proporción simple, obteniendo el porcentaje de raciones que pertenece a cada territorio. En otras palabras, el valor de alguna UT es la suma de los valores porcentuales de cada comuna que la conforma. Una UT es más atractiva si es que posee más raciones, debido a las economías de escala existentes en la operación. La Tabla 4 muestra un ejemplo del cálculo del atractivo de cada UT según el número de raciones.

UT	Cantidad de raciones	Atractivo de raciones
A	20,000	40.0%
B	30,000	60.0%
Total	50,000	

Tabla 4. Cálculo del atractivo de cada UT según el de número de raciones.

■ Número de Colegios

Al tratarse de un criterio inversamente proporcional no es posible que puedan sumarse los valores porcentuales individuales de las comunas para obtener el valor de la UT. La manera de resolver este problema es determinar el número de colegios que tiene cada UT formada y determinar un índice que consiste en dividir el número de colegios total de la región por el número de colegios de la UT. Por ejemplo, el valor de 2,22 de la UT A del ejemplo se obtiene al realizar 200/90. Con estos índices normalizados al 100 % se obtiene el atractivo para cada UT según el número de colegios (Ver Tabla 5).

UT	Cantidad de colegios	Índice	Atractivo de colegios
A	90	2.22	55.0%
B	110	1.82	45.0%
Total	200		

Tabla 5. Cálculo del atractivo de cada UT según el de número de colegios.

■ Superficie

La forma de obtener el valor es determinando el número de kilómetros cuadrados que posee cada UT formada. Como mientras menor es el área, más atractiva es la unidad, lo que se calcula es un índice que consiste en dividir el área total de la región por el área de la UT. Por ejemplo, el valor de 2,50 de la UT A del ejemplo se obtiene al realizar 20.000/8.000. Luego, como se muestra en la Tabla 6, se obtuvo la proporción de este resultado para todos los territorios, normalizando los índices calculados al 100 %.

■ Accesibilidad

Para obtener el peso de una UT según este criterio se procedió primero a calcular el índice de accesibilidad dividiendo para cada UT el número de colegios de fácil acceso por el número de colegios de difícil acceso (si no hay ningún colegio de éstos últimos en la UT se consideró que el número

UT	Cantidad de km ²	Índice	Atractivo de superficie
A	8,000	2.50	60.0%
B	12,000	1.67	40.0%
Total	20,000		

Tabla 6. Cálculo del atractivo de cada UT según su superficie.

de colegios de difícil acceso es 0,5 para evitar dividir por 0). Por ejemplo, el índice de accesibilidad de 21,50 de la UT A del ejemplo se obtiene al realizar 86/4. Luego de esto se obtiene el atractivo de accesibilidad de cada UT normalizando estos índices al 100 %. Sin embargo, puesto que este criterio es muy sensible a la cantidad de colegios de difícil acceso (por la poca cantidad), el resultado podría entregar datos distorsionados, en donde un único colegio adicional de difícil acceso puede pesar mucho al obtener el atractivo de accesibilidad de la UT. Esto se resolvió, como lo muestra la Tabla 7, suavizando el valor final de la siguiente manera: se pondera el valor obtenido por un 50 %, mientras que el otro 50 % corresponde al promedio que debiera tener cada unidad de la región en este ítem (para dos UT, un 50 %). Por ejemplo, el atractivo de accesibilidad "suavizado" (que es el valor final que usaremos para este criterio) de la UT A del ejemplo se calcula haciendo $0,5 * 65,7 + 0,5 * 50 = 57,9$.

UT	Cantidad de colegios de fácil acceso	Cantidad de colegios de difícil acceso	Índice de accesibilidad	Atractivo de accesibilidad	Atractivo de accesibilidad suavizado
A	86	4	21.50	65.7%	57.9%
B	101	9	11.22	34.3%	42.1%
Total	187	13			

Tabla 7. Cálculo del atractivo de cada UT según la accesibilidad a los colegios.

Resumiendo, en esta metodología los valores son obtenidos conociendo la configuración total de la región, posibilitando la comparación de un territorio con el resto de la región. Un cambio de alguna comuna varía el puntaje de todas las UT de la región, pues el valor final de alguna de éstas depende del valor del resto.

UT factibles, formadas por algoritmo	Atractivo de raciones	Atractivo de colegios	Atractivo de superficie	Atractivo de accesibilidad	*	Raciones	38.1%
A	40.0%	55.0%	60.0%	57.9%		Colegios	34.2%
B	60.0%	45.0%	40.0%	42.1%		Superficie	16.9%
						Accesibilidad	10.8%

UT factibles, formadas por algoritmo	Atractivo de UT
A	50.4%
B	49.6%

Tabla 8. Cálculo final del atractivo de cada UT utilizado en esta metodología.

Finalmente en la Tabla 8, para obtener la importancia relativa de cada

UT, se multiplica la matriz formada por los vectores de preferencia para cada UT de cada criterio (2x4) por la matriz de pesos de criterios (4x1), obteniéndose la matriz con los resultados finales para cada UT (2x1).

3.2. Descripción de la heurística

Entre las características más relevantes de la heurística de búsqueda local figura la toma de decisiones en función de la información que está disponible en cada iteración, pero la desventaja que tiene este método es que obviamente no garantiza alcanzar la solución óptima.

Diremos que una solución es vecina de otra si la nueva configuración se obtiene de la anterior por medio de pasar una comuna de una UT a otra (manteniendo constante el numero de UT de la región), o eliminar una UT integrando sus comunas a UT contiguas (disminuyendo en uno el número de UT), o transformar una UT en dos nuevas (aumentando en uno el número de UT).

La heurística procede iterativamente de la siguiente manera:

1. Se tiene un conjunto inicial de Unidades Territoriales predefinido en cada región, cada una formada por comunas. Esta configuración regional corresponde a la solución original que tiene JUNAEB hasta antes de la realización de este trabajo.
2. Se pasa a una solución vecina de la solución actual con el objetivo de “mejorar” la configuración existente
3. Se verifica que cada UT en la nueva configuración cumpla las restricciones de mínimo y máximo número de raciones. Si no se cumplen las condiciones, entonces se vuelve un paso atrás y no se considera el cambio de configuración hecho.
4. Se evalúa la desviación estándar. Si ésta disminuye entonces se mantiene el cambio; por el contrario, si aumenta, el cambio de configuración no es realizado.
5. El proceso termina cuando ninguna solución vecina puede disminuir la desviación estándar.

La desviación estándar (DE) es una medida de dispersión que mide cuánto tienden a alejarse los valores de cada UT con respecto al promedio de la región. Una desviación estándar grande indica que los valores están lejos de la media, y una desviación pequeña indica que los puntajes están agrupados cercanos a la media.

$$DE = \sqrt{\frac{n_r \sum x_j^2 - (\sum x_j)^2}{n_r^2}}$$

n_r = cantidad de UT de la región r y x_j = puntaje de la UT_j .

En la Figura 3 se muestra un ejemplo concreto (I Región) de cómo opera el algoritmo, junto con la desviación estándar y atractivo de las UT en cada iteración.

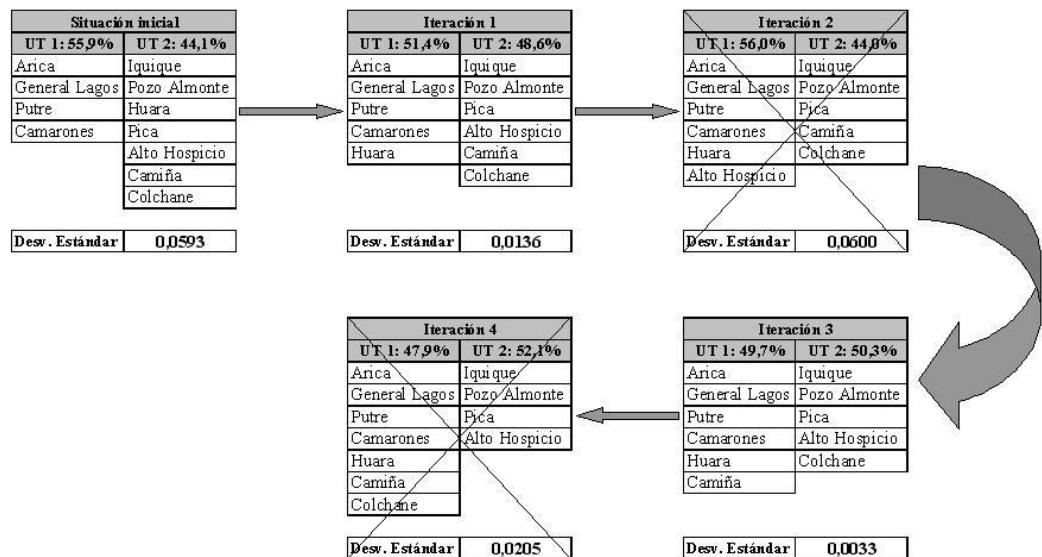


Figura 3. Ejemplo del funcionamiento de la heurística.

La iteración 2 no es realizada debido a que empeora la función objetivo, por lo que se vuelve a la iteración 1 y se realiza un intercambio de otra comuna. La iteración 4 tampoco es realizada porque la desviación estándar aumenta. Como no puede mejorarse la solución de la iteración 3 con un nuevo intercambio, el algoritmo finaliza.

Cabe señalar que aquellas comunas con más de 10.000 raciones en cualquier región son particionadas (y entonces la heurística trabajará con estas sub-comunas como si fueran comunas), de manera de repartir más homogéneamente dichas raciones con el objetivo de lograr UT con atractivos más similares. La partición en estos casos se hace en la menor cantidad de partes iguales necesaria para que cada sub-comuna tenga menos de 10.000 raciones (por ejemplo, una comuna de 24.000 raciones será partida en 3 sub-comunas de 8.000 raciones cada una).

La heurística fue programada en el lenguaje Java y el código fue compilado usando la versión de Java Standard Edition Development Kit (Java SE JDK) 1.5.0.06. El computador utilizado fue un Pentium 4, con un procesador intel

dual core de 1.8 gigahertz y memoria RAM de 4 gigabytes.

Si bien para mejorar una UT, se empeora otra, el resultado esperado es que mejore la eficiencia de las empresas y de JUNAEB en forma global, pues se estimarán más certeramente los costos asociados y habrá un control más estricto sobre los concesionarios. Una estimación de costos más exacta nace de la idea de que cada criterio usado tiene un costo con el cual puede relacionarse. Por ejemplo, la cantidad de raciones con el costo de materia prima y la superficie con el costo de transporte. Estos costos están intrínsecamente considerados en este estudio debido a que al comparar los criterios de a pares en el AHP, se están tomando en cuenta estos aspectos. La constitución de territorios de atractivo parejo significa un interesante resultado para las empresas, dado que facilita el cálculo de los costos en las distintas ofertas que presenten. Esto se debe a que la presencia de UT homogéneas por región permite a las empresas hacer sus cálculos de costos para una UT cualquiera y extenderlo en forma casi directa al resto de las UT de la región, a diferencia de lo que sucedía anteriormente en que los costos podían variar fuertemente de UT en UT, aún dentro de la misma región.

El estudio también hace un aporte desde lo social, ya que al equiparar las UT se espera que los servicios sean mucho más equitativos, lo que implica que no habrá niños que reciban una peor comida que otros. Obtener UT atractivas y parejas para un concesionario permitirá entregar un servicio no discriminatorio y con estándares de calidad aceptados por JUNAEB.

4. Modelo de Programación Entera

Otra metodología que se desarrolla para resolver el problema de la homogeneización de las UT es un modelo exacto de programación lineal entera. Para ello generaremos todas las posibles UT que podrían constituirse (que llamaremos *clusters*), formadas por comunas o particiones de comunas contiguas, con la condición de que cada una de ellas tenga a lo menos 15.000 raciones y a lo más 40.000. Luego definiremos un puntaje de atractivo para cada cluster, usando ideas similares a las desarrolladas en la Sección 3.1. Por último, mediante un modelo de programación lineal entera, seleccionaremos para cada región un conjunto de *clusters* que sean una partición de la región, minimizando la diferencia entre el cluster más atractivo y el menos atractivo. Estos *clusters* seleccionados constituirán la propuesta de partición en UT de la región.

4.1. Algoritmo de enumeración de *clusters*

Este algoritmo utiliza ideas contempladas en el método ARM (Area Restriction Model), usado en problemas forestales (ver [2,4]). En ARM se desarrolla una planificación de cosechas, respetando un máximo de área contigua que puede ser cosechada en un período. El diseño de este problema forestal contempla una región, la cual está dividida en bloques (los *clusters* de nuestro trabajo), los cuales, a su vez, están divididos en celdas (las comunas o particiones de comunas en este análisis). Éstas constituyen la unidad básica del problema y pueden ser cosechadas en un solo período (Murray y Weintraub [4]) o en diversos períodos (Goycoolea et al [2]). En este estudio, se trata de un problema de un solo período: la nueva configuración se hace una vez y se usa por un tiempo medianamente largo.

El problema es resuelto enumerando a priori todos los *clusters* que sean factibles en cuanto a contigüidad de comunas y número de raciones. El algoritmo de elaboración de *clusters* está basado en el usado en [3], que fue diseñado para un problema forestal buscando construir bloques de terrenos contiguos sujeto a un área máxima. En el presente estudio se hace una analogía a ese problema, manteniendo la contigüidad, pero la condición de área es reemplazada por un rango de raciones.

El número máximo de raciones que puede asignarse a cada unidad es de 40.000, así como el mínimo es de 15.000. Además se determinó que las celdas (comunas o particiones de comunas) pueden tener hasta 10.000 raciones. Por ello, las comunas que presentan un número superior a ese valor son divididas en partes iguales hasta que cada una de éstas tenga menos que esa cifra. Como el máximo de raciones que puede tener una celda es de 10 mil, entonces ninguna UT puede contener a una sola de ellas, puesto que el mínimo requerido es de 15 mil raciones, las cuales pueden ser alcanzadas con al menos dos. Cada celda debe ser asignada exactamente a una sola UT, la cual debe estar conformada sólo por terrenos contiguos, es decir, la región que queda conformada debe ser conexa en el sentido de que uno puede moverse de una celda a otra cualquiera de la misma UT sin salirse de la región.

El algoritmo empieza definiendo todos los posibles *clusters* compuestos de una celda, ubicándolos en el conjunto $S[1]$. Al empezar la $k+1$ -ésima iteración, el conjunto $S[k]$ contendrá todos los *clusters* factibles (teniendo en cuenta los rangos mínimo y máximo de raciones) de k celdas contiguas. Entonces, el conjunto $S[k+1]$ es construido luego de sumar de todas las formas posibles una celda contigua a cada miembro de los *clusters* de $S[k]$. La recursión finaliza cuando no quedan *clusters* factibles para formar o se alcanza el número máximo de iteraciones fijado.

El número de iteraciones determina el número máximo de celdas que puede tener un *cluster*. Para determinar este número se decidió darle cierta holgura

al resultado de la heurística. Por ejemplo, si en el resultado de la Primera Región, según la heurística, la UT con más celdas tenía 8 de ellas, entonces aquí fijábamos el máximo número de iteraciones en 10.

Los *inputs* usados por el algoritmo para cada región son:

1. Las celdas que forman parte de la región.
2. El número máximo de celdas que pueden formar parte de un mismo territorio, es decir el número máximo de iteraciones que puede hacer el algoritmo.
3. Una matriz de contigüidad espacial, la cual es una matriz binaria que indica las celdas contiguas a otras.
4. Un vector de raciones por cada celda. De esta forma es posible obtener la cantidad de raciones que cada *cluster* formado a priori obtiene.

Región	Celdas	Iteraciones	Clusters formados (algoritmo)	UT seleccionadas (modelo)
I	14	10	3.143	2
II	11	9	347	2
III	11	9	701	2
IV	19	10	17.875	3
V	34	7	17.072	7
VI	31	7	51.788	5
VII	32	7	55.731	6
VIII	45	7	46.874	10
IX	35	6	19.597	7
X	34	7	24.979	8

Tabla 9. Información sobre el algoritmo enumerador de cluster.

De la información mostrada en la Tabla 9, las celdas y los *clusters* formados son *inputs* para el modelo de programación entera (recordar que los *clusters* de sólo una celda no son incorporados como datos al modelo porque se sabe que no son factibles). A éstos deben sumarse los puntajes de cada *cluster*, los cuales son explicados en la Sección 4.2.

Cabe señalar que la cantidad de *clusters* formados no sólo depende del número de celdas de la región y de la cantidad máxima de celdas que pueden pertenecer a una UT. Existen otros dos aspectos que se encuentran implícitamente presentes. El primero es la cantidad de adyacencias existentes, en otras palabras, la cantidad de 1's que contiene la matriz de adyacencia; el segundo aspecto es la distribución geográfica que tienen las celdas, es decir, su ubicación dentro de la región. Una mayor cantidad de 1's en la matriz, por sí sola, no asegura que se forme una mayor cantidad de territorios, puesto que si bien es cierto que en la mayoría de los casos esta situación ocurre, también puede

pasar que la región tenga una silueta que separe intrínsecamente a las celdas en grupos y que entre éstos haya poca conexión. Un ejemplo de esta situación es lo que ocurre en la V, VIII y X regiones, en donde a pesar de haber más celdas y la misma cantidad de iteraciones que en la VII región, se obtienen menos *clusters*.

La programación del algoritmo, al igual que la heurística de búsqueda local, fue realizada en el lenguaje Java y el código fue compilado usando la versión de Java Standard Edition Development Kit (Java SE JDK) 1.5.0.06. Los programas que fueron implementados para el desarrollo del algoritmo utilizan también el mismo hardware que el usado en la heurística.

4.2. Pesos de cada UT por criterio utilizados en el modelo de PE

Notar que en este caso al calcular el valor de un *cluster* según cada criterio no se conoce la composición del resto de la región. De este modo, la forma de obtener los puntajes en esta metodología usa la misma base que en la heurística, sin embargo, los valores son calibrados, de manera que para una misma UT los atractivos obtenidos en la heurística sean similares que en el modelo de programación entera. Para ello, se toma como base el valor de alguna UT en la heurística para mantener aquí invariante ese valor, y obtener el resto por medio de una proporcionalidad directa. Tomaremos como ejemplo el mismo que fue utilizado en la Sección 3.1. En este caso, las UT A y B deben tener el mismo atractivo para cada criterio, mientras que las restantes deben tener atractivos proporcionales.

■ Número de Raciones

La Tabla 10 muestra los valores obtenidos para cada *cluster* según este criterio, calibrados para obtener un factor 40 de atractivo en la UT A y un factor 60 de atractivo en la UT B (que eran los porcentajes calculados al obtener los valores en la otra metodología). Notar que los valores de 40 y 60 son proporcionales a 20.000 y 30.000. El resto de los valores para los otros clusters que se muestran en este ejemplo se calcula manteniendo las proporciones de acuerdo a la cantidad de raciones de cada uno de ellos. Observar que por ejemplo *clusters* C y D, E y F, o G y H forman también diferentes particiones de la región.

■ Número de Colegios

La Tabla 11 muestra los valores obtenidos para cada *cluster* según este criterio, calibrados para obtener un factor de atractivo 55 en la UT A (que era el porcentaje calculado al obtener los valores en la otra metodología). El resto de los valores se calcula manteniendo las proporciones de acuerdo a los índices generados por la división entre el número

Clusters factibles formados por el algoritmo	Cantidad de raciones	Atractivo de raciones calibrado
A	20,000	40.0
B	30,000	60.0
C	32,500	65.0
D	17,500	35.0
E	25,000	50.0
F	25,000	50.0
G	35,000	70.0
H	15,000	30.0

Tabla 10. Cálculo del atractivo de cada cluster según el de número de raciones.

de colegios total de la región y el número de colegios del *cluster*. Por ejemplo, el atractivo 49,5 del *cluster* C se obtiene haciendo $55*2/2,22$. Recordemos que los clusters C y D, E y F, o G y H forman también una partición de la región, tal cual sucedía para la UT A y la UT B, pero sin embargo las sumas de los atractivos en los *clusters* de la misma partición no son necesariamente iguales a 100. Algunos experimentos numéricos y los resultados finales de este estudio nos muestran que esta calibración (y la efectuada para los criterios de superficie y accesibilidad) permite calcular de manera realista el atractivo final de cada *cluster* en esta metodología.

Clusters factibles formados por el algoritmo	Cantidad de colegios	Índice	Atractivo de colegios calibrado
A	90	2.22	55.0
B	110	1.82	45.0
C	100	2.00	49.5
D	100	2.00	49.5
E	80	2.50	61.9
F	120	1.67	41.3
G	110	1.82	45.0
H	90	2.22	55.0

Tabla 11. Cálculo del atractivo de cada cluster según el de número de colegios.

■ Superficie

La Tabla 12 muestra los valores obtenidos para cada *cluster* según este criterio, calibrados para obtener un factor de atractivo 60 en la UT A (que era el porcentaje calculado al obtener los valores en la otra metodología). El resto de los valores se calcula manteniendo las proporciones de acuerdo a los índices generados por la división entre el número

de km^2 total de la región y el número de km^2 del *cluster*. Por ejemplo, el atractivo 43,6 del *cluster* C se obtiene haciendo $60*1,82/2,50$.

Clusters factibles formados por el algoritmo	Cantidad de superficie	Índice	Atractivo de superficie calibrado
A	8,000	2.50	60.0
B	12,000	1.67	40.0
C	11,000	1.82	43.6
D	9,000	2.22	53.3
E	10,000	2.00	48.0
F	10,000	2.00	48.0
G	12,000	1.67	40.0
H	8,000	2.50	60.0

Tabla 12. Cálculo del atractivo de cada *cluster* según la superficie.

■ Accesibilidad

La Tabla 13 muestra los valores obtenidos para cada *cluster* según este criterio, calibrados y suavizados para obtener un factor de atractivo final igual a 57,9 en la UT A (que era el porcentaje calculado al obtener los valores en la otra metodología). Para ello se calculan primero los atractivos de acceso de cada UT tomando como base el valor de 65,7 de la UT A. El resto de los valores para obtener el atractivo de acceso calibrado se calcula manteniendo las proporciones de acuerdo al índice de accesibilidad (índice que se calcula haciendo la división entre el número de colegios de fácil acceso y el número de colegios de difícil acceso del *cluster*). Por ejemplo, el atractivo 30,9 del *cluster* C se obtiene haciendo $65,7*10,11/21,50$. Finalmente, para obtener los atractivos finales (ya suavizados) de cada *cluster*, recordemos que en este criterio el índice influye en un 50 % en el valor final, mientras que el otro 50 % corresponde al promedio que debiera tener cada *cluster* en este ítem. Como en esta metodología no se conoce a priori la composición total de la región, para obtener el promedio lo que se considera es el número de celdas que tiene el *cluster* y ese valor se lo divide por el total de celdas de la región.

En resumen, en esta metodología se calcula el atractivo de cada *cluster* sin conocer lo que ocurre en el resto de la región. De esta manera, los valores deben ser calibrados para obtener puntajes que mantengan la proporcionalidad entre los *clusters*.

Al igual que lo realizado para la heurística, para obtener el atractivo de un *cluster* en esta metodología se multiplica la matriz formada por los vectores de preferencia para cada *cluster* de cada criterio (8x4), por la matriz de pesos de criterios (4x1), obteniéndose la matriz con los atractivos para cada *cluster*, lo que se muestra en la Tabla 14 para el ejemplo desarrollado.

Clusters factibles formados por el algoritmo	Cantidad de colegios de fácil acceso	Cantidad de colegios de difícil acceso	Índice de accesibilidad	Atractivo de acceso calibrado	Atractivo de acceso calibrado y suavizado
A	86	4	21.50	65.7	57.9
B	101	9	11.22	34.3	42.1
C	91	9	10.11	30.9	40.4
D	96	4	24.00	73.3	61.7
E	76	4	19.00	58.1	54.0
F	113	7	16.14	49.3	49.7
G	102	8	12.75	39.0	44.5
H	86	4	21.50	65.7	57.9

Tabla 13. Cálculo del atractivo de cada cluster según la accesibilidad.

Clusters factibles, formados por algoritmo	Atractivo de raciones	Atractivo de colegios	Atractivo de superficie	Atractivo de accesibilidad
A	40.0	55.0	60.0	57.9
B	60.0	45.0	40.0	42.1
C	65.0	49.5	43.6	40.4
D	35.0	49.5	53.3	61.7
E	50.0	61.9	48.0	54.0
F	50.0	41.3	48.0	49.7
G	70.0	45.0	40.0	44.5
H	30.0	55.0	60.0	57.9

*

Raciones	38.1%
Colegios	34.2%
Superficie	16.9%
Accesibilidad	10.8%

UT factibles, formadas por algoritmo	Atractivo de UT
A	50.4
B	49.6
C	53.4
D	45.9
E	54.2
F	46.6
G	53.6
H	46.6

Tabla 14. Cálculo del atractivo final de cada cluster.

4.3. Formulación del modelo

La distribución de celdas para la confección de *clusters* es un problema de un alto grado de complejidad, principalmente por la gran cantidad de combinaciones existentes entre las celdas que pueden formar los *clusters*. Estas posibles UT serán el *input* principal de nuestro modelo de programación entera. Como este tipo de modelos suelen ser de difícil resolución (recordemos que en términos de complejidad computacional la programación entera es un problema *NP-Hard*), es importante buscar alternativas para la función objetivo que no agreguen muchas variables al problema.

En un principio se intentó homologar la desviación estándar minimizada en la heurística, utilizando variables auxiliares y teniendo como objetivo la minimización de la suma de los valores absolutos de las diferencias entre puntajes de las unidades seleccionadas. Este objetivo obligaba a tener una gran cantidad de nuevas variables, que agrandarían fuertemente los tiempos de resolución del problema.

Por ello se prefirió la siguiente alternativa. Los *clusters* obtenidos a priori son usados como variables de decisión en el modelo de programación lineal entera. Además, cada *cluster* posee un valor, que se calcula tal cual fue explicado en la Sección 4.2. La función objetivo del modelo busca minimizar la diferencia entre el mayor y menor puntaje de una UT dentro de una región. Este objetivo constituye una aproximación a la no linealidad de la desviación

estándar. El modelo de optimización es una herramienta que permite comparar la bondad de los resultados arrojados por la heurística, dado que ambas funciones objetivo (minimizar la desviación estándar y minimizar la brecha entre la “mejor” y la “peor” UT) tienen características similares. Además, el modelo también fue utilizado para generar buenas soluciones iniciales desde donde empezar a aplicar la heurística (ver Sección 5). La formulación de este problema se muestra a continuación.

■ Variables

a = puntaje mínimo de una UT dentro de una región

b = puntaje máximo de una UT dentro de una región

$$y_j = \begin{cases} 1 & \text{si el } cluster j \text{ es seleccionado como UT} \\ 0 & \sim \end{cases}$$

■ Parámetros

v_j = indica el puntaje del *cluster* j, obtenido con los cuatro criterios utilizados.

Ω_i = conjunto de *clusters* que contienen a la celda i.

P= promedio de puntajes de las UT seleccionadas en una región.

■ Función objetivo

$$\min(b - a) \quad (1)$$

■ Restricciones

$$\sum_{j \in \Omega_i} y_j = 1 \quad \forall i \quad (2)$$

$$a \leq v_j * y_j + (1 - y_j) * P \leq b \quad \forall j \quad (3)$$

$$y_j \in \{0, 1\} \quad \forall j \quad (4)$$

$$a, b \geq 0 \quad (5)$$

La formulación tiene como fin (ecuación (1)) la obtención de UT relativamente homogéneas, minimizando la brecha entre el mayor (b) y el menor (a) puntaje en cada región. La ecuación (2) dice que de todos los *clusters* que contienen a una celda dada se debe elegir como UT a exactamente uno. Las inecuaciones de (3) establecen el rango en el cual los puntajes de las UT deben oscilar. El parámetro P corresponde al puntaje promedio de las UT que se seleccionarán. Dado que a priori no se conoce este promedio (pero se puede estimar), se utilizaron entre tres y cinco valores distintos de P para cada región, eligiéndose finalmente el que otorgaba una mejor función objetivo. La restricción (4) indica que y_j es una variable binaria, mientras que (5) indica que a y b deben ser números reales positivos.

Esta formulación del modelo fue programada en el software GAMS 22.1 y resuelta por CPLEX 10.0. El modelo fue ejecutado bajo el mismo hardware que la heurística y el algoritmo de enumeración de *clusters*.

5. Método de Programación Entera + Heurística

En la última metodología empleada, a la solución entregada por el modelo de programación entera que minimiza la brecha entre la “mejor” y la “peor” UT, se le aplicó la heurística, buscando disminuir la desviación estándar. La idea de esta nueva metodología tiene origen en el resultado de la heurística y el modelo de programación entera para la Primera Región. El hecho de que el modelo dé una desviación estándar un poco menor, aún en un caso de pocas comunas, da la idea de que la solución inicial en la que está configurada la región es importante para un mejor resultado de la heurística.

La heurística y el modelo de programación entera son dos buenos complementos a la hora de obtener un resultado que satisfaga objetivos similares. Así, usando el modelo de programación entera para minimizar la diferencia entre el mayor y menor puntaje dentro de una región, es posible usar ese resultado como el paso 0 de la heurística. La solución entregada por el modelo de programación entera no debiera sufrir un gran cambio cuando se le aplica la heurística en las regiones donde es posible mejorar la desviación estándar, debido a que la solución inicial es óptima para minimizar la brecha entre la “mejor” y la “peor” UT, y a la relación directa entre ambos objetivos.

6. Resultados

En este capítulo se muestran los resultados obtenidos con las metodologías antes descritas. Se evaluó para cada una de las regiones de Chile, excepto la XI,

XII y Metropolitana, las mejores configuraciones obtenidas con los tres métodos. Mientras la XI y la XII no fueron analizadas por no tener las suficientes raciones para formar al menos dos UT, la Metropolitana se omitió porque dada la gran cantidad de comunas adjudicadas a empresas Pymes que presentaba, se pretendía mantenerla con la configuración utilizada hasta 2007. Además, en este estudio se analizaron las regiones instituidas en Chile hasta 2006, por lo que las actuales regiones de Los Ríos y de Los Lagos son estudiadas en forma conjunta, al igual que Arica-Parinacota y Tarapacá.

El estudio fue utilizado por JUNAEB con el objetivo de modificar su actual segregación de territorios para sus licitaciones, utilizando el 100 % de las alternativas postuladas en este trabajo. Se seleccionó para cada región la solución de menor desviación estándar obtenida por la heurística, que fue la primer metodología que se terminó.

En la licitación de 2007, cuyo contrato comenzó a ejecutarse a partir del mes de Marzo de 2008, ya se realizó la segregación territorial basada en el estudio para las regiones de Arica/Parinacota, Tarapacá, Antofagasta, del Maule, de Los Ríos, de Los Lagos y 1/3 de la Región Metropolitana. Con esto, se ha licitado un tercio de los servicios entregados por JUNAEB en todo el país. En las próximas licitaciones, las de 2008, para iniciarse en 2009 y la del año 2009, para iniciarse en 2010, se realizará la segregación territorial utilizando también el estudio aquí desarrollado.

En algunas regiones el cambio entre la distribución original de las comunas y la entregada por cualquiera de los tres métodos resultó ser muy grande, debido a una sustancial heterogeneidad que presentaban las regiones antes de este trabajo, mientras que en otras el cambio fue pequeño, pues las UT ya eran bastante homogéneas. La nomenclatura usada es:

- **O = Configuración original.** Se calcula el valor de la desviación estándar
- **H = Configuración elaborada con la heurística.** Objetivo: “Minimizar la desviación estándar entre las UT de una región”
- **M = Configuración elaborada con modelo de programación entera.** Objetivo: “Minimizar la diferencia entre el mayor y el menor puntaje de las UT de una región”
- **M+H = Configuración elaborada con la heurística, tomando como solución inicial el resultado entregado por el modelo de programación entera.** Objetivo: “Minimizar la desviación estándar”

Los resultados para cada método, a priori, deberían ser los mismos o muy parecidos para las regiones con pocas comunas. Sin embargo para las regiones

con muchas comunas pueden dar resultados variados debido a la alta cantidad de combinaciones que pueden darse, la cual aumenta exponencialmente.

En la Tabla 15 se presenta una comparación del valor de la desviación estándar (multiplicada por 1000) en el puntaje de las UT en las nuevas configuraciones obtenidas por cada uno de los métodos aquí desarrollados y en la configuración hasta 2007. El porcentaje de mejora se obtiene a través de la división entre el módulo de la diferencia de ambas desviaciones y la desviación más alta.

Región	Desviación estándar*1000				Variaciones desviación estándar					
	Original	Heurística	Modelo	M+H	O v/s H	O v/s M	O v/s M+H	H v/s M	H v/s M+H	M v/s M+H
I	59.2606	3.3235	2.9002	2.9002	94.39%	95.11%	95.11%	12.74%	12.74%	0.00%
II	1.7656	1.7656	1.7656	1.7656	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
III	7.3759	3.0138	3.0138	3.0138	59.14%	59.14%	59.14%	0.00%	0.00%	0.00%
IV	59.3588	17.4444	8.7488	6.4244	70.61%	85.26%	89.18%	49.85%	63.17%	26.57%
V	14.6982	7.2983	4.2298	2.9939	50.35%	71.22%	79.63%	42.04%	58.98%	29.22%
VI	32.6641	4.9151	4.1361	4.1361	84.95%	87.34%	87.34%	15.85%	15.85%	0.00%
VII	40.5145	8.1358	5.3422	2.0553	79.92%	86.81%	94.93%	34.34%	74.74%	61.53%
VIII	24.3784	8.2864	7.8469	5.5500	66.01%	67.81%	77.23%	5.30%	33.02%	29.27%
IX	39.0247	8.1090	0.8062	0.8062	79.22%	97.93%	97.93%	90.06%	90.06%	0.00%
X	10.3572	8.6694	8.7741	8.6694	16.30%	15.28%	16.30%	1.19%	0.00%	1.19%

Tabla 15. Comparación de la desviación estándar multiplicada por 1000 entre los distintos métodos.

Para realizar las comparaciones se calcularon los atractivos de las UT en cada solución final, utilizando en todos los casos la primera de las metodologías (desarrollada en la Sección 3.1).

Comparando los resultados de la desviación estándar en cada caso puede observarse en general una notoria variación entre la distribución empleada por JUNAEB hasta el año 2007 y los resultados de los tres métodos. Además, el modelo de programación entera a pesar de no buscar el objetivo de minimizar la desviación estándar, obtiene resultados un poco mejores que la heurística en ese sentido, exceptuando la X región, aunque en este caso se trata de un aumento de sólo un 1,19 %. Esto se debe a que en la mayoría de los casos existe una correlación entre ambos objetivos; es decir que si los resultados del modelo son los mejores minimizando la diferencia entre el mayor y el menor puntaje, también permiten disminuir la desviación estándar. De este modo es razonable pensar que la solución óptima del modelo de programación entera pueda ser usada como solución inicial de la heurística, cuando se busca minimizar la desviación estándar. Las mejores soluciones las entrega la metodología del modelo de programación entera + heurística, pues comienza con una solución inicial que es óptima para el objetivo de minimizar la diferencia, la cual es mejorada por la heurística en la búsqueda de minimizar la desviación estándar.

En la Tabla 16 se muestra la brecha entre los valores máximo y mínimo de cada UT en los resultados finales obtenidos en cada método y en la configuración inicial. El porcentaje de mejora se obtiene a través de la división entre

el módulo de la diferencia de ambas brechas y la brecha más alta.

Para realizar las comparaciones se calcularon los atractivos de las UT en cada solución final, utilizando en todos los casos la segunda de las metodologías (desarrollada en la Sección 4.2).

El mejor resultado lo entrega obviamente el modelo de programación entera, dado que tiene como objetivo minimizar dicha diferencia (notar que el modelo alcanza el óptimo en todos los casos). La diferencia es mayor con los otros dos métodos en aquellas regiones con una gran cantidad de comunas y, por consiguiente, mayor cantidad de combinaciones para formar *clusters*. Además, es interesante notar que, a pesar de que la heurística y el modelo de programación entera + heurística tienen como objetivo minimizar la desviación estándar, el segundo siempre presenta mejores resultados que el primero, también para el caso de disminuir la brecha entre la mejor y la peor UT de cada región.

Región	Diferencias (máx - mín)				Variaciones entre métodos					
	Original	Heurística	M+H	Modelo	O v/s H	O v/s M+H	O v/s M	H v/s M+H	H v/s M	M+H v/s M
I	12.70	0.55	0.04	0.04	95.67%	99.69%	99.69%	92.73%	92.73%	0.00%
II	0.20	0.20	0.20	0.20	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
III	2.55	1.15	1.15	1.15	54.90%	54.90%	54.90%	0.00%	0.00%	0.00%
IV	46.49	4.63	2.36	1.58	90.04%	94.92%	96.60%	49.03%	65.87%	33.05%
V	12.62	2.84	1.53	1.27	77.50%	87.88%	89.94%	46.13%	55.28%	16.99%
VI	40.84	2.62	0.54	0.54	93.58%	98.68%	98.68%	79.39%	79.39%	0.00%
VII	63.48	2.45	1.73	0.17	96.14%	97.27%	99.73%	29.39%	93.06%	90.17%
VIII	72.86	4.86	2.93	2.49	93.33%	95.98%	96.58%	39.71%	48.77%	15.02%
IX	47.09	2.72	0.15	0.15	94.22%	99.68%	99.68%	94.49%	94.49%	0.00%
X	3.81	2.52	2.52	2.26	33.86%	33.86%	40.68%	0.00%	10.32%	10.32%

Tabla 16. Comparación de diferencias entre el mayor y menor puntaje de cada UT para cada metodología.

Es importante ver a través de estos resultados que con cualquier punto de vista analizado se obtienen valores más homogéneos en las UT de cada región. Esta situación muestra una cierta robustez de los métodos empleados en este trabajo.

Como se mencionó anteriormente, existía en el caso del modelo de programación entera una complicación adicional para calcular el puntaje de cada *cluster* puesto que el mismo se obtiene sin conocer la distribución del resto de la región (a diferencia de lo que ocurre con la heurística). Sin embargo, los buenos resultados obtenidos por el modelo en todas las regiones dan la idea de que la calibración de estos valores ha sido adecuada.

Lo que se mantiene inalterable en todas las metodologías es el número de UT óptima, como se explicita en la Tabla 17. Era de esperar una muy baja variación en este valor, dado que una estimación del mismo es fácilmente obtenido al calcular el número promedio de raciones por unidad (a los efectos de este cálculo ayuda mucho conocer las cotas inferiores y superiores de raciones por UT). Notar que el número total de UT se redujo en 25, un 18%

del total de 136.

Región	UT antes del estudio	UT después del estudio
I	2	2
II	2	2
III	2	2
IV	5	3
V	10	7
VI	9	5
VII	10	6
VIII	17	10
IX	9	7
X	11	8
Total	77	52

Tabla 17. Cantidad de UT antes y después del estudio.

Los tiempos de resolución variaron considerablemente de acuerdo a la metodología empleada. En la heurística los tiempos de corrida son de sólo unos milisegundos para las regiones con pocas comunas, mientras que para las regiones con más comunas los tiempos pueden aumentar hasta sólo un par de segundos. En este aspecto influye la configuración original que tenía JUNAEB, la cantidad de comunas y la distribución que presentan las comunas dentro de la región, pues no es lo mismo que las comunas con muchas raciones estén aglomeradas a que se encuentren dispersas alrededor de toda la región.

El algoritmo de elaboración de *clusters* es también rápido. En la mayoría de las regiones la demora en la conformación de los *clusters* factibles no llegó al minuto. Sólo en las regiones VI, VII y VIII el tiempo llegó a los 20 minutos. Este tiempo depende de la cantidad de celdas, del número máximo de celdas que pueden pertenecer a un *cluster* y de la cantidad de adyacencias en la región.

Los tiempos del modelo de programación entera no superan la hora de corrida, excepto en la VIII Región, donde puede llegar hasta las 5 horas, debido a la cantidad de celdas y a la gran cantidad de combinaciones existentes. Las regiones I, II y III no superan los 60 segundos en su resolución. Los aspectos más importantes que influyen en los tiempos de resolución del modelo es la cantidad de *clusters* que satisfacen las condiciones del algoritmo y la distribución de las celdas atractivas.

Finalmente cuando el resultado del modelo es usado como solución inicial de la heurística, los tiempos no varían demasiado con respecto al segundo método, pues los tiempos de corrida de la heurística, como ya se ha dicho, son de unos pocos segundos.

7. Conclusiones

En la mayoría de las regiones la nueva distribución de las UT entre los distintos análisis es similar y puede decirse que la heurística es un método muy rápido para obtener resultados satisfactorios.

La menor desviación estándar se logra con el método modelo + heurística, pues se inicia con una solución óptima para uno de los objetivos que es variada para mejorar el otro objetivo. En contraparte, la menor diferencia entre el mayor y menor puntaje se logra con el modelo de programación entera, el cual entrega la solución óptima para el objetivo buscado en todos los casos.

La configuración utilizada hasta 2007 mostraba importantes problemas en la VI y VII Región dado que las comunas muy atractivas eran adjudicadas íntegramente a una UT. Esto ocasionaba que las firmas se interesaran mucho más por esas UT en desmedro del resto de la región. Las propuestas actuales dividen estas comunas y se las asigna en más de una UT, de tal forma de que haya varios territorios con una importante cantidad de raciones para una compañía que postula a la licitación.

Con la propuesta de configuración territorial realizada disminuye el riesgo de bancarrota de alguna empresa que pertenece a este proceso, pues ya no existirán UT poco atractivas.

Según JUNAEB el objetivo principal del problema fue cumplido a cabalidad, pues la nueva segregación utilizada permitió mayor interés en la adjudicación de nuevas zonas para las empresas oferentes y, además, éstas han podido hacer más competitivas sus ofertas. Esto mismo permitirá a JUNAEB controlar de mejor manera la gestión y operación de los concesionarios dentro de una región.

Cabe destacar que la confección de una nueva configuración que cumple con los objetivos de JUNAEB y los requerimientos de las empresas fue posible gracias a la definición de los criterios principales que determinan el atractivo de una UT, a la cuantificación de los mismos y al uso de técnicas de programación matemática que en base a estos criterios permitieron conseguir configuraciones con atractivos homogéneos. Claramente, un enfoque manual como el empleado para hacer la configuración que regía hasta 2007 no hubiera permitido alcanzar este tipo de resultados.

Un tercio de la nueva configuración propuesta fue utilizada por JUNAEB, licitando durante 2007 las siguientes regiones distribuidas en UT por la heurística: Arica-Parinacota, Tarapacá, Antofagasta, del Maule, de Los Ríos, de Los Lagos y un tercio de la Región Metropolitana (región no analizada en el presente trabajo). Durante los próximos dos años se seguirá usando el resto del

resultado entregado.

Agradecimientos: Al Instituto de Ciencias Milenio "Sistemas Complejos de Ingeniería" por el apoyo brindado para la concreción de este estudio. A las autoridades de JUANEBC, especialmente a Amalia Cornejo y Lizette Vega, por toda su colaboración. A Flavia Bonomo y Pablo Rey, por sus sugerencias que permitieron mejorar este trabajo. A Nicolás Maluenda y Ronald Poillot, por su ayuda en la implementación de la heurística de búsqueda local. El segundo autor está parcialmente financiado por el Fondecyt 1080286. El tercer autor está parcialmente financiado por el Fondecyt 1085188.

Referencias

- [1] Caro F., Shirabe T., Guignard M. and Weintraub A., School redistricting: embedding GIS tools with Integer Programming, Journal of the Operational Research Society (2004) 55, 836-849.
- [2] Goycoolea M., Murray A., Barahona F., Epstein R. and Weintraub, A., Harvest Scheduling Subject to Maximum Area Restrictions: Exploring Exact Approaches, Operations Research (2005) 53(3), 490-500.
- [3] Goycoolea, M., Murray, A., Vielma, J. and Weintraub, A., Evaluating Approaches for Solving the Area Restriction Model in Harvest Scheduling,, European Journal of Operational Research (2007), submitted.
- [4] Murray, A. and Weintraub, A., Scale and Unit Specification Influences in Harvest Scheduling with Maximum Area Restriction,. Forest Science (2002) 48(4), 779-789.
- [5] Saaty, T., The Analytic Hierarchy Process, McGraw Hill (1980).
- [6] Taylor R., Vasu M. and Causby J., Integrated planning for school and community: the case of Johnston County, North Carolina, Interfaces (1999) 29 (1), 67-89.
- [7] Zamorano, G., Métodos Cuantitativos para Configuración de Unidades Territoriales Aplicado a JUNAEB. Tesis (Magíster en Gestión de Operaciones e Ingeniería Civil Industrial). Santiago, Chile. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, 2008. 150 págs.

METODOLOGÍAS BASADAS EN VAR PARA EL ANÁLISIS DE RIESGO EN PROYECTOS DE EMBALSES

PROSPER LAMOTHE*

EDUARDO CONTRERAS**

Resumen

Esta investigación aplica el modelo de Value at Risk al mercado agrícola en Chile, en particular a proyectos de concesión de obras de riego. En estos proyectos, la viabilidad del negocio para el concesionario de la obra de riego, depende de la rentabilidad del negocio de los agricultores (clientes del concesionario), y también de los riesgos de la ejecución de las obras. La rentabilidad del negocio agrícola, depende a su vez de una serie de factores de riesgo, entre los que destacan los riesgos de mercado (de precios de los productos) y los riesgos climáticos. Como principal conclusión, se valida la hipótesis de que es posible estimar rentabilidades y riesgos para el concesionario de un proyecto de embalse, mediante una metodología como la desarrollada.

Palabras Clave: Análisis de riesgo - VaR - Evaluación de Proyectos.

* Académico. Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid.

** Académico. Facultad de Ciencias Físicas y Matemáticas. Departamento de Ingeniería Industrial. Universidad de Chile.

1. Introducción

La aplicación de VaR a proyectos de riego es relevante, ya que en el estado chileno, que inició en la década de los noventa un sistema de concesión de obras públicas a empresas privadas (fundamentalmente carreteras, puentes, aeropuertos y puertos), se ha analizado la posibilidad de extender este sistema la construcción de obras de riesgo.

En este sentido, el aporte de la investigación se relaciona con la novedad de aplicar técnicas de VaR a portafolios de commodities agrícolas (cuya rentabilidad determina la del embalse). La correcta estimación de los riesgos de este tipo de proyectos, contribuirá a facilitar la toma de decisiones de las empresas privadas participantes, evitando costos asociados a una inadecuada valoración de los riesgos.

2. Marco Teórico de la Investigación

2.1. Análisis probabilístico

Consiste en calcular estimadores de tendencia central y de dispersión del VPN de un proyecto de inversión a través de su función de distribución de probabilidades. Para un proyecto es posible calcular el VPN esperado y la desviación estándar de dicho VPN (Blanck y Tarquin, 1998). En este enfoque juega un papel importante el Teorema del Límite Central: Si una variable aleatoria X puede ser expresada como la suma de n variables aleatorias independientes, entonces para un "n grande" la variable aleatoria X sigue aproximadamente una distribución normal (Jorion, 2004).

El VPN es en efecto una variable aleatoria que es igual a la suma (ponderada) de $n+1$ variables aleatorias: los flujos de caja. Por lo que con las funciones de distribución de los flujos de caja se puede obtener el comportamiento probabilístico del VPN.

2.2. Nuevos desarrollos: el Value at Risk (VaR)

La metodología VaR (Longerstaey J. & Zangar P, 1995) ha sido ampliamente difundida, aplicada a distintos tipos de activos, fundamentalmente financieros. Lo que el VaR mide es la exposición al riesgo para un cierto nivel de confianza, es decir, el monto máximo que se podría perder para ese nivel de confianza, en cierto intervalo de tiempo.

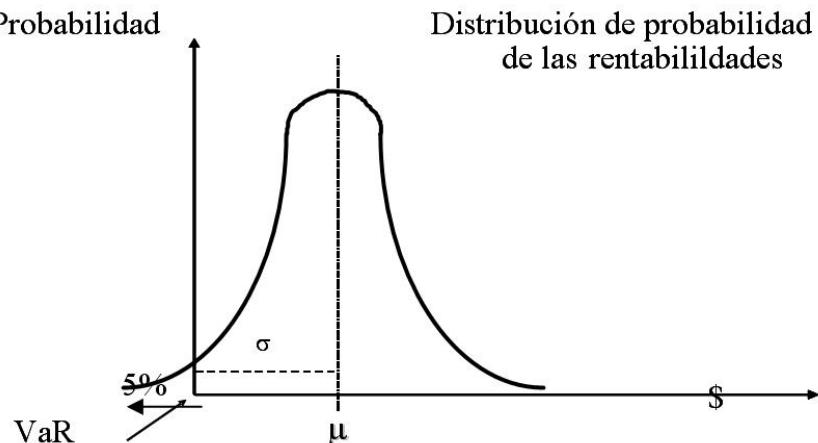


Figura 1: Value at Risk (VaR)

La métrica se ha establecido en el mercado como una medida clave del riesgo, desde que el Comité de Basilea estableció que los bancos deberían ser capaces de cubrir pérdidas en sus portafolios sobre un horizonte de 10 días un 99 % del tiempo.

Relación entre el VaR y las opciones reales

Existe una complementariedad teórica y práctica entre el VaR y el enfoque de opciones reales, este último busca valorar el aporte de las decisiones futuras de los gestores de un proyecto, que introducen mejoras en la gestión dependiendo de la evolución de las variables riesgosas (una síntesis de estas “flexibilidades” se puede analizar en Contreras y Fernández, 2003). Ambos consideran que las variables riesgosas evolucionan a futuro según algún proceso estocástico. Más aún, el modelamiento, por ejemplo de los precios, que se realice para calcular el VaR, debe ser el mismo modelamiento que se utilice para calcular el valor de las opciones asociadas a la evolución futura de dichos precios (Alessi, 2005).

Desde el punto de vista práctico, tanto en el cálculo del VaR como en el de las opciones reales se puede utilizar un modelo paramétrico, asumiendo normalidad de las variables (lo que conduce al modelo de Black y Sholes en el caso de opciones y al modelo RiskMetrics en el caso de VaR), o se pueden estimar las variabilidades de las variables mediante simulación de Monte Carlo, es decir, mediante un modelo no paramétrico (Embrechts. 2003). En síntesis, hay mucho en común entre los dos enfoques, los que se pueden considerar complementarios, sólo que apuntan a objetivos distintos: en el caso del VaR se busca medir en nivel de riesgo de un activo financiero o un proyecto, mientras que con las opciones reales se busca medir el valor de las flexibilidades que se

generan en razón de la existencia de aquellos niveles de riesgo.

Dado que el objetivo de este trabajo es medir riesgos en proyectos de embalses, la metodología propuesta y la aplicación están orientadas al cálculo del VaR de los proyectos, pero se incluye además una aplicación de opciones reales.

3. Fundamentos del VaR

El VaR o Value at Risk, es hoy una de las herramientas más utilizadas por los gerentes de Riesgo en la medición de riesgo de mercado. Su medición tiene fundamentos estadísticos y el estándar de la industria es calcular el VaR con un nivel de significancia de 5 %. Esto significa que solamente el 5 % de las veces, o 1 de 20 veces, es decir una vez al mes con datos diarios, el retorno del portafolio caerá más de lo que señala el VaR, en relación del retorno esperado.

Analíticamente el VaR (Jorion, 2004), se define por el límite superior de la integral de la función de los rendimientos esperados $r(s)$.

$$\begin{aligned} E[r] - VaR \\ \int_{-\infty}^{-VaR} r(s)ds = \alpha \end{aligned} \quad (1)$$

Usualmente se asume que el valor esperado de los rendimientos es cero, $E[r]=0$, con lo cual la solución a la expresión anterior se transforma en:

$$\int_{-\infty}^{-VaR} r(s)ds = \alpha$$

Una representación alternativa consiste en estimar el VaR como:

$$VaR = \alpha\sqrt{\sigma^2\Delta t} \quad (2)$$

donde α es el factor que define el área de perdida de los rendimientos, σ^2 la varianza de los rendimientos, y Δt el horizonte de tiempo para el cual se calculará el factor de riesgo VaR. Si el valor medio es distinto de cero, el VaR se puede medir de forma relativa al valor medio (Dowd, 1999).

La ecuación (2) representa el VaR del retorno de un activo. Si se busca el VaR del valor, en lugar del retorno, debe tenerse en cuenta que el valor V es igual al número de unidades de activo N multiplicado por el precio del mismo ($V=N*P$). Se puede verificar entonces que el cambio de valor $(V_1 - V_0)/V_0 = \alpha \sqrt{\sigma^2 t}$, de donde se desprende que la ecuación (2) de los retornos, para el caso del VaR del valor de una activo, se transforma en:

$$VaR = V_0\alpha\sqrt{\sigma^2 t} \quad (3)$$

El riesgo total medido como VaR, corresponde entonces a la pérdida potencial máxima que puede ocurrir a causa de un evento riesgoso, con un determinado intervalo de confianza, y en un determinado horizonte de tiempo. Analíticamente, lo que se quiere es encontrar un VaR tal que:

$$P = \text{Prob}\{\Delta V \leq \text{VaR}\} \quad (4)$$

Donde V representa una función de valor tal como el valor presente de un activo.

3.1. Cálculo del VaR paramétrico de una cartera

Un portafolio (o un proyecto), está compuesto por posiciones activas y pasivas que son afectadas por factores de riesgo. Estas posiciones tienen un Value at Risk que depende de la volatilidad de los factores de riesgo que las afectan.

En teoría, los portafolios están construidos de tal manera, que la relación entre los instrumentos que lo componen hacen que el riesgo agregado sea menor que la suma de los riesgos individuales (Markowitz, 1952). Por lo anterior, la contribución de cada posición al riesgo del portafolio depende de la relación entre los diferentes factores de riesgo que lo componen. Esta relación entre factores de riesgo puede ser medida a través del coeficiente de correlación (ρ).

Una vez determinada la correlación entre los diferentes factores de riesgo que componen el portafolio, se puede proceder a calcular el VaR diversificado. El Value at Risk de un portafolio puede estimarse como:

$$(\text{VaR}_{\text{portafolio}})^2 = \sum_{i=1}^n (\text{VaR}_i)^2 \rho_i^2 + \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{VaR}_i \text{VaR}_j \rho_{ij} \quad (5)$$

donde:

n: número de factores de riesgo.

VaR_i : Value atRisk del factor de riesgo i.

ρ_{ij} : Correlación entre el factor de riesgo i y el factor de riesgo j.

$\text{VaR}_{\text{portafolio}}$: Value at Risk del portafolio que equivale a $\sqrt{(\text{VaR}_{\text{portafolio}})^2}$

Esta suma no sólo contiene los VaR individuales, los cuales están representados en el primer sumando, sino también todos los distintos productos cruzados representados por el segundo término. Para efectos de la aplicación a un embalse, para llegar a aplicar la ecuación (5) primero se han de calcular los VaR individuales de los factores de riesgo con la ecuación (2). En el caso particular de los precios de productos agrícolas, el Value at Risk se determina a partir de la desviación estándar de los precios de los productos, el valor

asociado a las ventas de dicho productos (equivalente al peso de una activo dentro de una cartera), y considerando un valor a de la distribución normal correspondiente a un 95 % de confianza.

3.2. Otras metodologías de cálculo de VaR: Simulación

Hay tres metodologías de cálculo de VaR que han adquirido mayor difusión:

La Simulación Histórica: Esta es una simulación simple, que requiere relativamente pocos supuestos respecto a las distribuciones estadísticas de los factores de riesgo. Para el caso de un activo consiste en ordenar los valores históricos de una variable (por ejemplo el retorno de una acción) y ordenarlos de menor a mayor para construir un histograma. Dicho histograma permite encontrar el valor que deja a la izquierda un 5 % de los casos (o un 1 % según el nivel de confianza que se desee para el VaR). Este método resulta difícil de aplicar al VPN de un proyecto, ya que se requerirían series históricas de VPN (que no existen) o calcular VPN para cada set de valores de las variables riesgosas (precios de los productos agrícolas), pero asumiendo que cada set de valores de las variables se mantiene constante en el horizonte de evaluación.

El Método Paramétrico de varianzas / covarianzas presentado en 3.1, está basado en el supuesto de que los factores de riesgo subyacentes en el mercado siguen una distribución normal mutivariada.

El método de Simulación de Monte Carlo: tiene similitudes con la simulación histórica. En la de Monte Carlo los valores son obtenidos simulando mediante la generación aleatoria de valores de acuerdo a alguna función de distribución. En el caso de varios factores de riesgos, este método se basa en información suficiente sobre la distribución conjunta de estas variables. Entonces generando valores de acuerdo a esta distribución conjunta podemos generar un gran número de escenarios, y para cada uno de ellos calcular un VPN, de forma que un número elevado de escenarios nos permite obtener una buena aproximación a la distribución del VPN. El menor quintil (1 %, 5 % según el nivel de confianza) de esa distribución será la aproximación al VaR (Glasserman, P., Heidelberger P. and Shahabuddin.P. 2000).

Compación entre las tres metodologías

Los métodos difieren en su potencial para capturar riesgos de opciones, o de activos semejantes a opciones (como las opciones reales). Difieren también en su facilidad de implementación, su facilidad para presentar y explicar resultados, la flexibilidad para analizar efectos de cambios de supuestos y por

último en la confiabilidad de los resultados. El mejor método podrá ser seleccionado entonces dependiendo de cuál de los aspectos anteriores importa más (Feria, 2005) (Jorion, 2004).

Los métodos de simulación trabajan bien independientemente de la presencia de opciones en los portafolios, porque en estos métodos se recalculan los valores de mercado para cada "jugada" de los factores básicos de mercado, incluyendo los valores implícitos de las opciones. En contraste, el método de varianzas-covarianzas (método paramétrico), funciona bien para instrumentos y portfolios con limitados contenidos de opciones, son menos capaces de capturar los riesgos de este tipo que los métodos de simulación. (Best, 1999). Tomando en cuenta que los proyectos de embalses tienen implícitas opciones (de cambio de cultivos por ejemplo), se considera más apropiada la simulación de Monte Carlo.

La simulación histórica y el método paramétrico de varianzas - covarianzas son más fáciles de implementar que la simulación de Monte Carlo. Los dos primeros requieren solo de datos históricos de las variables riesgosas. Ambos se pueden desarrollar sin mayor dificultad en un software utilitario básico como Excel.

La simulación de Monte Carlo requiere, además de los datos históricos, de un modelamiento computacional, de preferencia con algún software ad hoc para simulación o una programación para esos efectos (Matlab, Visual Basic, Macros sobre Excel, Risk, u otro). Adicionalmente Monte Carlo requiere simular la distribución conjunta de las variables riesgosas, para lo cual se requiere hacer descomposición de Cholesky lo cual nuevamente requiere softwares ad hoc o programación.

Un problema único de la simulación de Monte Carlo, es el hecho de que el modelador del sistema puede elegir la distribución de probabilidad de los factores de mercado, esta flexibilidad le permite hacer una mala elección.

Como síntesis del análisis comparativo anterior; se privilegiará por el mayor potencial de la simulación de Monte Carlo para capturar el valor de las opciones. Por tratarse de una investigación, son menos relevantes las desventajas que pueda tener el método en términos de su mayor dificultad de implementación y la mayor dificultad para presentar y explicar los resultados.

4. Pasos previos: análisis de sensibilidad y metodología

Como paso preliminar, se realizó un análisis de sensibilidad para identificar las variables relevantes a analizar. Estos análisis se realizaron para tres proyectos y se llegó a los siguientes resultados.

Proyecto Ancoa: La elasticidad a los precios es entre 5 y 6 veces mayor que la de la inversión, y entre 3 y 4 veces mayor que la elasticidad a la tasa de incorporación a riego.

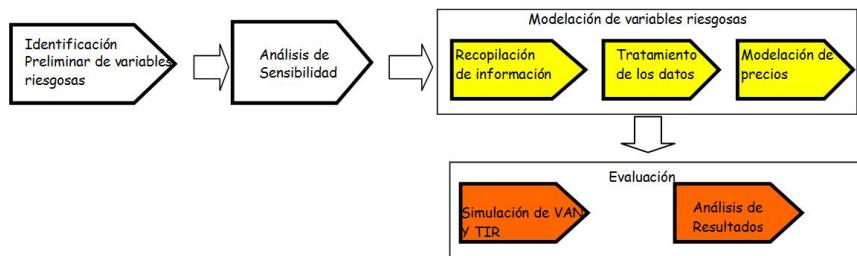
Proyecto Laja-Diguillín: La elasticidad a los precios es entre 6 y 10 veces mayor que la de la inversión, y aproximadamente 2 veces la de la tasa de incorporación a riego.

Proyecto Victoria de Vilcún: La elasticidad a los precios es aproximadamente 6 veces mayor que la de la inversión, y aproximadamente 4 la de la tasa de incorporación a riego.

Todos los resultados consistentemente muestran la importancia del conjunto de precios de productos agrícolas como variables de riesgo. No obstante lo anterior, se modelaron también los riesgos relativos a la inversión, pero teniendo presente que las variables más importantes de modelar son los precios¹. El riesgo climático es relevante para la situación sin proyecto, pero su eliminación es justamente la fuente de beneficios y de disposición a pagar por parte de los agricultores (en la situación con proyecto), y por tanto el diferencial de riesgos queda capturado en el modelamiento agro-económico de la metodología de evaluación vigente. El riesgo climático no cubierto por el diseño del embalse, se puede considerar con la inclusión de eventos aleatorios de "Saltos" de Poisson (esta forma de modelar se incluye en el capítulo 7).

La figura describe la metodología propuesta para el método no paramétrico.

Figura 2: Diagrama de flujo del procedimiento de evaluación bajo incertidumbre.



Fuente: Elaboración propia

¹Cabe señalar que, por falta de datos, no se pudo sensibilizar con respecto a los rendimientos de los terrenos (los que a su vez dependen del clima).

En el punto siguiente se describe el modelamiento de los precios. Para el análisis de los puntos de recopilación de información y tratamiento de datos, ver Contreras (2008).

5. Aplicación de la metodología propuesta

5.1. Procesos de precios

Para modelar precios de acciones y commodities en los métodos de simulación se supone (entre otros modelos) que los precios siguen un Proceso de Wiener generalizado (Vilarino, 2001) (Jorion, 2004) (Feria, 2005), esto es, que tiene un drift constante μ y una varianza constante σ^2 , en la práctica se supone que el drift y la varianza sean una proporción del precio S , de aquí se puede mostrar entonces el modelo de comportamiento de precios conocido como Movimiento Browniano Geométrico a través de la siguiente ecuación diferencial estocástica:

$$\frac{dS(t)}{dt} = \mu[t, S(t)]dt + \sigma[t, S(t)]dW_t \quad (6)$$

donde $dS(t)$ representa el cambio en los precios, S , en un intervalo pequeño de tiempo, dt , $dW_t = \xi \sqrt{t}$ (proceso de Gauss-Wiener, con $\xi \sim N(0,1)$), luego la ecuación anterior está normalmente distribuida con media μdt y desviación estándar $\sigma \sqrt{t}$. En otras palabras,

$$\frac{dS(t)}{dt} \rightarrow \phi(\mu dt, \sigma \sqrt{dt}) \quad (7)$$

Ahora supongamos que el precio S , sigue un proceso de Ito (Lamothe y Pérez Somalo, 2004), se demuestra (elevando ambos términos de (6) al cuadrado y en función de la tabla de multiplicación aplicable a las integrales estocásticas) que:

$$dG = \left(\frac{dG}{dS} \mu S + \frac{dG}{dt} + \frac{1}{2} \frac{d^2G}{dS^2} S^2 \sigma^2 \right) dt + \frac{dG}{dS} \sigma S dW \quad (8)$$

Se deriva entonces el proceso seguido por $\ln(S)$ (el logaritmo natural del precio) utilizando el lema de Ito. Sea $G = \ln S$, luego

$$\frac{dG}{dS} = \frac{1}{S} \quad \frac{d^2G}{dS^2} = -\frac{1}{S^2} \quad \frac{dG}{dt} = 0$$

Reemplazando los términos en el proceso seguido por G , se llega a:

$$dG = (\mu - \sigma^2/2)dt + \sigma dW \quad (9)$$

Reemplazando en la última ecuación dG por $d\ln(S)$, se llega a:

$$d\ln(S) = (\mu - \sigma^2/2)dt + \sigma dW \quad (10)$$

Por lo tanto para modelar los precios, se debiera considerar esta distribución, es decir:

$$\ln S_T \sim \phi[\ln S + (\mu - \frac{\sigma^2}{2})T, \sigma\sqrt{T}] \quad (11),$$

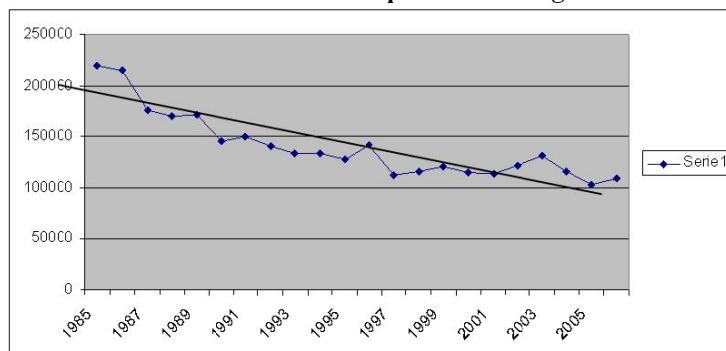
donde $\phi(m, s)$ denota una distribución normal con media m y desviación estándar σ . Esta expresión se puede discretizar (Vilariño, 2001) como:

$$\ln S_T = [\ln S + (\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}\xi] \quad (12)$$

En lo que respecta a modelos de volatilidad, Brooks y Persans (2002) indican que Alexander y Leigh (1997) examinaron, modelos GARCH, modelos EWMA y modelos de promedios de pesos similares (sin darle mayor peso al reciente), encontrando como resultado que éste último resulta superior a los otros dos, este método de promedio simple es el utilizado en este trabajo.

Se elige el modelo de movimiento browniano geométrico, principalmente por su mayor simplicidad, y en segundo lugar porque los precios históricos mayoritariamente siguen una tendencia. A modo de ejemplo, a continuación se presentan la serie de precios del trigo. La existencia de estas tendencias sugiere que este modelamiento es mejor que el de modelos sin drift como el de Ornstein y Uhlenbeck (descrito en Ergashev, 2002) de reversión a la media sin tendencia o las series de tiempo de ruido blanco”, aunque se podrían considerar también modelos de reversión a la media con tendencia, si bien estos últimos son más complejos de implementar².

Gráfico 1: Serie de precios del Trigo



Fuente: Elaboración propia a partir de datos de ODEPA

²Se agradece al profesor José Miguel Cruz, del Departamento de Ingeniería Industrial, por sus aportes en este análisis de procesos con o sin reversión a la media en el caso de precios de productos agrícolas.

Existen desde luego modelos alternativos para precios de comodites. Por ejemplo Engel y Valdés (2002) concluyen que modelos AR (1) simples y de camino aleatorio, son más precisos para el mediano plazo (hasta 5 años, no hacen el análisis para periodos superiores a 5 años) en el caso de precios de commodities, que modelos más complejos (modelos AR(1) con coeficientes estocásticos estimados mediante el filtro de Kalman). Un modelo de movimiento browniano, es un modelo de camino aleatorio continuo en logaritmos (Metcalf y Hasset, 1995), y en ese sentido los resultados de Engel y Valdés sugieren la conveniencia de usar un modelo como el de la ecuación (12) en el corto plazo, pero como se señala más adelante, en el largo plazo se espera que los precios converjan hacia algún precio de equilibrio (el costo marginal si el mercado es competitivo). Una alternativa se puede ver en Lamothe y Otero (2003), que aplican un proceso de Metcalf & Hasset (en lugar del movimiento browniano geométrico) con tendencia y reversión a la media.

Más allá de este análisis teórico, se realizaron análisis estadísticos a las series de precios. Antes de realizar los test de raíz unitaria, es conveniente inspeccionar los correlogramas de las series. Los cuadros construidos (Contreras, 2008) permitieron apreciar que en todos los casos, las autocorrelaciones caen muy lentamente en el tiempo. Por otra parte, para varias de las series, el primer coeficiente de autocorrelación parcial es alto (mayor que 0.7), siendo los coeficientes restantes cercanos a cero, ello es indicativo de la potencial presencia de una raíz unitaria, lo que a su vez probaría que los precios siguen un camino aleatorio con drift y que por ende los retornos logarítmicos siguen un proceso browniano geométrico.

El próximo paso consiste en llevar a cabo test de raíces unitarias para cada serie (Fernández, 2007), a fin de verificar las conjeturas anteriores. De acuerdo a tests de Dickey-Fuller aumentados (ADF), en todos los casos (exceptuando la serie de la palta), no rechazamos H_0 (presencia de una raíz unitaria) al 99 por ciento de confianza. Se aplicó el test de Dickey y Fuller aumentado (ADF) mediante el uso de E-views. Por lo tanto, se puede afirmar que empíricamente no se observa la reversión, aún cuando teóricamente debería estar en el largo plazo.

En efecto, los movimientos brownianos tienden a alejarse de sus valores iniciales, lo cual es apropiado para algunas variables como los precios de los activos especulativos pero, en otros casos, aunque los valores de corto plazo se muevan aleatoriamente, en largo plazo tienden a ajustarse hacia promedios históricos o a precios de equilibrio (Villalpando, 2007). En estos casos, se dice que la variable sigue un proceso de reversión hacia la media.

Como ya se ha dicho, por simplicidad se trabajará con modelos de procesos brownianos, aún cuando en la práctica las volatilidades (con reversión a la media), debieran ser menores que las estimadas con el modelo seleccionado. Para ilustrar lo anterior, se han realizado dos simulaciones de precios para

productos del valle de La Ligua, se seleccionaron cinco productos: palta, maíz, porotos granados, naranjas, mandarinas y limones. Con los precios de estos productos se construyó una cartera (en la que cada producto “pesa” según su porcentaje de contribución al margen) y se simularon precios futuros a 20 años para finalmente calcular un valor presente de los precios simulados con procesos brownianos y con procesos con reversión. Para este último se estimaron coeficientes de reversión a la media mediante regresiones de mínimos cuadrados, y se consideró que los precios de equilibrio de largo plazo eran los precios promedio históricos de cada producto.

Tabla 1: Comparación procesos estocásticos con y sin reversión a la media.

Nombre de hoja de resultados	Reversión a la Media
Número de variables simuladas	5
Número de Iteraciones	100
Media	586.993927
Desviación Estandar	167.0213067
Varianza	27896.11689
Valor Mínimo	\$ 435.21
Valor Máximo	\$ 1,385.81

Nombre de hoja de resultados	Proceso Browniano
Número de variables simuladas	5
Número de Iteraciones	100
Media	340.652515
Desviación Estandar	251.5801141
Varianza	63292.5538
Valor Mínimo	\$ 138.48
Valor Máximo	\$ 1,530.03

Fuente: Elaboración propia

En la tabla 1 se observa que efectivamente el proceso browniano estima una volatilidad mayor³. Debe tenerse en cuenta que esta diferencia no es necesariamente la que podría observarse en las aplicaciones a proyectos que se presentan más adelante, ya que en esos casos se simularán más productos, y estas simulaciones se aplicarán tanto a los precios de los productos en la situación con proyecto, como a los de la situación sin proyecto, además de que se consideran riesgos en otras variables.

En lo sucesivo se trabajará con procesos brownianos, tomando en cuenta: su mayor facilidad de implementación (y por ende de adopción por parte de los tomadores de decisiones), el respaldo empírico de los test de raíces unitarias, y el hecho de que se trabajará con valores presentes, indicador con el cual los flujos en los que se generan las mayores discrepancias entre ambos modelos, son los que pesan menos en valor actual, aminorando así la diferencia.

³Además estima una media menor, ya que considera la tendencias (todas negativas excepto la del limón), sin que estas reviertan a una media de largo plazo, es decir, sin que se frenen esas tendencias.

En un trabajo conjunto con la Dirección de Obras Hidráulicas del Ministerio de Obras Públicas, se hicieron simulaciones para las rentabilidades de varios proyectos de embalses: Ancoa, Laja, Diguillín, Victoria de Vilcún, Ligua y Petorca (estos dos últimos son los que se presentan en el punto siguiente). Adicionalmente se analizaron los proyectos Quino y Traiguén, para calibrar los modelos. En varios de estos proyectos se modeló además el sobre costo o sobre plazo de la inversión.

5.2. Aplicación de la metodología no paramétrica al proyecto Ligua - Petorca

Se desarrolló un programa en Visual Basic (Carmona, 2007), que tomaba como input las planillas de evaluación del modelo determinístico (Ayala y Cabrera, 2007). Con este programa se modelaron los precios para los cultivos. Una alternativa habría sido modelar directamente la variabilidad de los flujos de caja (la idea de Cash Flow at Risk, de Stein, Usher y otros, 2001), pero como a la fecha no existen proyectos concesionados de riego, no se cuenta con la información histórica necesaria.

Para estudiar el nivel de efectividad del modelamiento, y de su valor predictivo, se tomará como historia (input) 25 años y se estimarán los siguientes 5 años. Como se posee registro sobre los últimos 30 años, se calculan errores medios de los 5 últimos años comparando la serie de valores reales versus los estimados.

Para los resultados obtenidos con 10.000 iteraciones, se realizaron los Tests RMSE, MAE, MAPE y Thail Inequality Coefficient, comparando los modelos con y sin correlación, y la estimación determinística basada en el promedio simple de los 20 primeros años como pronóstico de los seis siguientes.

Tabla 2.- Test de pronósticos.

TEST DE PRONÓSTICOS		10.000 simulaciones		
Precios Cultivos		no correl	correl	Promedio (20 y)
RMSE		5660.6503	5623.1371	15404.9892
MAE		55.0804	54.9272	97.0281
MAPE		10.5020	10.4654	35.1131
Thail Inequality Coefficient		0.0525	0.0521	0.1224

Fuente: Elaboración propia

Se puede apreciar que los modelos con y sin correlación superan a la estimación determinística en todos los Tests (errores menores que la estimación determinística) siendo siempre el modelo con correlación algo más preciso que el sin correlación. El Thail Inequality Coefficient, mide el ajuste, de forma que el valor cero implica ajuste perfecto, y el valor uno implica un pésimo ajuste. Se validan así los modelos desarrollados, al menos en el horizonte de 5 años.

Modelo Estocástico para Ligua y Petorca

El modelo se construyó sobre la planilla elaborada por el equipo consultor de Ayala y Cabrera, en particular para todos los datos determinísticos no modelados (fundamentalmente costos) se asume que los valores de dichas planillas son correctos. De esa fuente se tomaron los costos de Inversión, control de heladas, costos de impulsión, costos de riego y gastos indirectos.

Con estos datos se obtuvo un VPN determinístico para Ligua de 15,921 millones y para Petorca de 36,782 millones. A priori esta evaluación debiese diferir de la que se haga mediante simulación, ya que las proyecciones de precios de productos agrícolas de Ayala y Cabrera no consideran ninguna tendencia de precios.

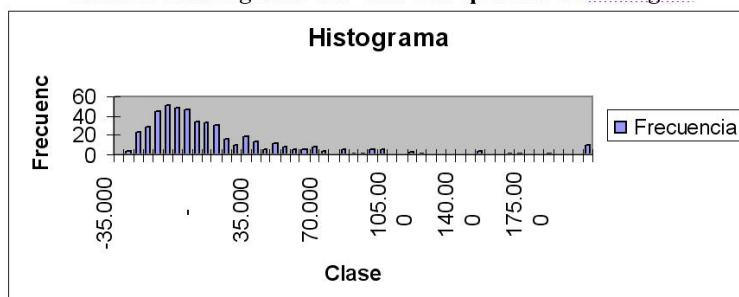
Para generar VaR por simulación, es importante mencionar que no todos los cultivos poseían precios históricos para el periodo de análisis. Por tanto, se seleccionaron los siguientes: chirimoya, mandarina, papaya, cebada, maíz, trigo, ají, choclo, poroto, porotos verdes, tomate, arveja, palta, lechuga, naranja, limón, lúcumo.

La tasa de descuento utilizada en la valoración, y en las sucesivas valoraciones en este trabajo, es la tasa libre de riesgo relevante para Chile (4,5 %). La determinación de esta tasa está tomada de las estimaciones de Cartes, Contreras y Cruz (2005)

Ahora bien, la simulación estática implicaría el supuesto poco realista de que se produce y se comercializan cultivos aún con márgenes negativos. Corrigiendo esto, es decir, no permitiendo márgenes negativos (lo que equivale a trabajar con un enfoque de valoración de opciones reales), se obtienen los resultados siguientes:

El valor esperado del VPN resultó \$19.610.000.000, y el VaR con 95 % de confianza es de -\$27.000.000.000, y la probabilidad de VPN negativo es de 39,8 %. Si se consideran los Earnings at Risk (EaR), es decir, el mejor 5 %, este da un valor de +110.000.000.000. La tabla de frecuencias transformada en histograma, se presenta a continuación:

Gráfico 2: Histograma del VPN con opciones de La Ligua.



Fuente: Elaboración propia

En el caso del proyecto Petorca, se obtuvo un VPN esperado de 26.430. Además del valor esperado se observa que el VaR con el 95 % de confianza es de (aproximadamente) -\$47.000.000.000, y que la probabilidad de obtener VPN negativo es de 49,8 %. El EaR es de +\$180.000.000.000.

Los procesos de precios por tener tendencias negativas disminuyen la rentabilidad en relación a la evaluación determinística, mientras que las opciones reales tienden a aumentar la rentabilidad, el efecto neto termina siendo en el caso de Petorca una disminución de rentabilidad esperada del modelo no paramétrico en relación al estudio determinístico, y en el caso de la Ligua se calcula una mayor rentabilidad en valor esperado respecto a la calculada bajo el enfoque determinístico.

Se pudo determinar la existencia de probabilidades de VPN negativo distintas de cero, lo cual constituye información valiosa para la toma de decisiones.

5.3. Estimación del VaR paramétrico para Ligua

En la estimación paramétrica que se presenta a continuación, a priori se espera algún sesgo negativo en cuanto a riesgo respecto al modelo no paramétrico, ya que el modelo paramétrico no captura las opciones. Por esa razón se espera que el VaR paramétrico sea mayor, dado que el VaR no paramétrico incluye flexibilidades que desplazan el VaR hacia la derecha de la distribución, haciendo asímétrico el histograma del VPN.

Para poder desarrollar el modelo paramétrico se requiere la matriz de correlaciones de los productos. Esta matriz (que denominaremos B) se presenta en la página siguiente. Para obtener los pesos de cada tipo de cultivo dentro del total de predios de la cuenca, se utilizaron las fichas de cultivos construidas a partir de las encuestas y el trabajo de terreno del estudio de Ayala y Cabrera. Los datos de volatilidades utilizados en el método paramétrico son los mismos que se usaron en el método no paramétrico. Estas volatilidades se presentan ya ponderadas por los pesos de cada producto en la cartera en la tabla 4 (le denominamos Vector A).

A continuación se calcula el vector C (producto de A por la matriz de correlaciones) y por último se multiplica este vector C por el vector A traspuesto (A') y se obtiene así el valor del riesgo de la cartera de productos agrícolas: 5,85 %. La desviación estándar (Riesgo) por tanto resulta ser la raíz del valor anterior, que arroja un resultado de 24,23 %.

Si se tratara de matrices de dos por dos, el modelo que se está construyendo sería el siguiente:

$$\text{Riesgo} = [w \cdot \sigma_1 \quad (1-w) \cdot \sigma_2] \cdot \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \cdot \begin{bmatrix} w \cdot \sigma_1 \\ (1-w) \cdot \sigma_2 \end{bmatrix} = w^2 \cdot \sigma_1^2 + (1-w)^2 \cdot \sigma_2^2 + 2 \cdot w \cdot (1-w) \cdot \sigma_1 \cdot \sigma_2 \cdot \rho \quad (13)$$

Donde la matriz B corresponde a la de correlaciones y el vector A es del de riesgos ponderados por pesos.

Tabla 3: Matriz de correlaciones

Matriz B: Correlaciones									
	Choclo	Arveja Verde	Tomate	Cebada	Chirimoyo	Mandarinas	Lechuga	Limon	
Choclo	100.00%	-18.71%	31.50%	1.78%	-43.59%	-16.71%	3.80%	-19.88%	
Arveja Verde	-18.71%	100.00%	-18.19%	10.67%	40.56%	-26.44%	19.24%	-3.53%	
Tomate	31.50%	-18.19%	100.00%	-10.76%	-55.70%	-12.81%	5.16%	-11.88%	
Cebada	1.78%	10.67%	-10.76%	100.00%	9.90%	-15.15%	11.30%	-0.03%	
Chirimoyo	-43.59%	40.56%	-55.70%	9.90%	100.00%	6.82%	8.06%	20.78%	
Mandarinas	-16.71%	-26.44%	-12.81%	-15.15%	6.82%	100.00%	-35.94%	10.46%	
Lechuga	3.80%	19.24%	51.6%	11.30%	8.06%	-35.94%	100.00%	-2.44%	
Limon	-19.88%	-3.53%	-11.88%	-0.03%	20.78%	10.46%	-2.44%	100.00%	
Lucuma	9.50%	-8.27%	11.75%	-0.05%	-12.92%	-1.73%	8.91%	7.62%	
Maiz	-3.13%	-2.01%	-6.33%	14.41%	-8.90%	-17.09%	15.77%	-3.87%	
Naranja	-5.70%	0.12%	-13.21%	-3.76%	8.74%	4.50%	-9.76%	2.29%	
Palta	-5.86%	7.85%	-6.33%	3.62%	5.33%	1.36%	-6.96%	-5.20%	
Papas	26.89%	-23.97%	32.12%	-0.65%	-43.31%	-1.21%	-2.88%	-4.74%	
Porotos Granados	49.87%	-16.49%	44.35%	-10.37%	-12.36%	-7.29%	-2.05%	-2.28%	
Porotos Verdes	35.87%	20.6%	47.52%	5.30%	-45.64%	-11.81%	0.96%	-25.81%	
Trigo	21.04%	-1.96%	14.76%	4.85%	-28.72%	-1.65%	4.14%	-10.80%	
Aji	27.12%	-17.67%	52.25%	-19.47%	-44.95%	-29.18%	2.06%	-8.42%	
Papaya	-4.62%	-21.74%	13.49%	-1.27%	5.30%	15.13%	-13.08%	0.02%	

Matriz B: Correlaciones										
	Lucuma	Maiz	Naranja	Palta	Papas	Porotos Granados	Porotos Verdes	Trigo	Aji	Papaya
Choclo	9.50%	-3.13%	-5.70%	-5.86%	26.89%	49.87%	35.87%	21.04%	27.12%	-4.62%
Arveja Verde	-8.27%	-2.01%	0.12%	7.85%	-23.97%	-16.49%	2.06%	-1.96%	-17.67%	-21.74%
Tomate	11.75%	-6.33%	-13.21%	-6.33%	32.12%	44.35%	47.52%	14.76%	52.25%	13.49%
Cebada	-0.05%	14.41%	-3.76%	3.62%	-0.65%	-10.37%	5.30%	4.85%	-19.47%	-1.27%
Chirimoyo	-12.92%	-8.90%	8.74%	5.33%	-43.31%	-12.36%	-45.64%	-28.72%	-44.95%	5.30%
Mandarinas	-1.73%	-17.09%	45.0%	1.36%	-1.21%	-7.29%	-11.81%	-1.65%	-29.18%	15.13%
Lechuga	8.91%	15.77%	-9.76%	-6.96%	-2.88%	-2.05%	0.96%	4.14%	2.06%	-13.08%
Limon	7.62%	-3.87%	2.29%	-5.20%	-4.74%	-2.28%	-25.81%	-10.80%	-8.42%	0.02%
Lucuma	100.00%	4.23%	-1.25%	-12.86%	-0.22%	12.45%	8.96%	8.83%	19.75%	-19.90%
Maiz	4.23%	100.00%	-0.17%	7.86%	5.52%	-14.48%	3.60%	9.82%	-1.52%	-14.97%
Naranja	-1.25%	-0.17%	100.00%	4.75%	9.04%	-7.65%	-1.39%	-4.75%	-9.04%	3.73%
Palta	-12.36%	7.86%	4.75%	100.00%	-5.00%	-14.80%	2.89%	0.28%	-6.57%	-5.94%
Papas	-0.22%	5.52%	9.04%	-5.00%	100.00%	41.33%	26.32%	8.24%	20.46%	6.43%
Porotos Granados	12.45%	-14.48%	-7.65%	-14.80%	41.33%	100.00%	58.92%	30.24%	22.06%	23.35%
Porotos Verdes	8.96%	3.60%	-1.39%	2.89%	26.32%	58.92%	100.00%	27.40%	44.99%	-13.79%
Trigo	8.83%	9.82%	-4.75%	0.28%	8.24%	30.24%	27.40%	100.00%	16.33%	-3.93%
Aji	19.75%	-1.52%	-9.04%	-6.57%	20.46%	22.06%	44.99%	16.33%	100.00%	-11.48%
Papaya	-19.90%	-14.97%	-14.97%	-5.94%	6.43%	23.35%	-13.79%	-3.93%	-11.48%	100.00%

Fuente: Elaboración propia a partir de series de precios de ODEPA.

Se deben multiplicar los pesos (obtenidos a partir de las estructuras de cultivos) por los riesgos individuales, de esta forma se obtiene el vector A:

Tabla 4: Riesgos ponderados por pesos en La Ligua.

Riesgos ponderados por pesos

Vector A	1.070%	1.059%	0.783%	0.273%	0.707%	1.806%	0.519%	2.141%	0.068%
----------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Vector A	0.418%	0.385%	24.091%	2.997%	2.639%	0.644%	0.256%	0.335%	1.067%
----------	--------	--------	---------	--------	--------	--------	--------	--------	--------

Fuente: Elaboración propia.

Se obtuvo un riesgo (varianza) de la cartera de productos agrícolas de 5,85 %. La desviación estándar por tanto resulta ser la raíz del valor anterior,

que arroja un resultado de 24,23 %. Finalmente, se calcula el VaR, fijando el valor esperado del VPN en 19.610 (ver Tabla 5), que es el valor obtenido mediante la simulación, para ese valor se despeja el valor actualizado de los beneficios netos (descontando la inversión actualizada) y sobre ese monto se aplica la variabilidad determinada para la cartera. Se obtuvo un VaR de \$141.028. El valor anterior corresponde al VaR relativo, es decir, la pérdida con respecto al valor esperado. En términos absolutos, el VaR es de -\$121.418.

Tabla 5: Cálculo del VaR en La Ligua

<i>Parámetro 5%</i>	-1.6449		5.00%
<i>Parámetro 95%</i>	1.6449		95.00%
<i>Rentabilidades</i>		<i>Resultados</i>	<i>VaR</i>
<i>Límite Alto</i>	54.28%	\$ 160,638.33	\$ 141,028
<i>Límite Medio</i>	14.42%	\$ 19,610.00	-
<i>Límite Bajo</i>	-25.43%	-\$ 121,418.33	\$ 141,028

Fuente: Elaboración propia.

Se obtiene un VaR considerablemente menor que el obtenido mediante el método no paramétrico (-\$27.000). La probabilidad de VPN negativo resultó 41 % (el percentil en el que el VPN es aproximadamente igual a cero), este valor se compara con una probabilidad de VPN negativo de 39,8 % obtenida mediante el método no paramétrico.

Como se anticipaba, la medida de riesgo es mayor con el método paramétrico. Este sesgo tiene que ver con la no valoración de las opciones reales en este modelo. No deben haber diferencias por las tendencias en los precios, ya que estas están implícitas en la valoración paramétrica por que se tomó como VPN esperado el que se obtuvo mediante el método no paramétrico. De hecho esto se hizo justamente para centrar el análisis comparativo en la medición de riesgos (y no en la medición de rentabilidades).

6. Calibración del modelo no paramétrico

Para validar el método, se aplicó el modelo a tres proyectos antiguos: Victoria de Vilcún, Quino y Traiguén. Esto permite simular los procesos de precios con la información disponible hasta el momento en que se evaluaron esos proyectos años atrás, y luego contrastar los VPNs esperados (ex ante) según la simulación del modelo, con el VPN real ex post que se obtiene de reevaluar el proyecto con los precios reales observados.

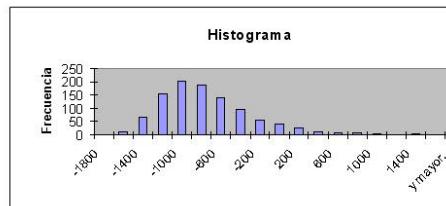
A continuación se presentan los resultados para Quino y Traiguén (la Tabla 9, incluye también Victoria de Vilcún). Se comparan los resultados de las evaluaciones de 1991 (Figueroa y otros, 1993), con las simulaciones del modelo y con los VAN ex post calculados con precios de los cultivos en el período 1991-

2006. Se trabajó con cifras en millones de \$ de 1991. Considerando los cambios de precios reales entre 1991 y 2006 (se toma como base 1992 para hacer los resultados comparables con los del modelo determinístico), el VPN ex post de Traiguén fue de \$430 millones y el de Quino fue de \$-1056 millones. Mientras que los resultados de las simulaciones fueron:

Tabla 6: Resultados simulación del VPN para Quino (millones de \$).

Media	-860
Desviación estándar	453
Mínimo	-1779
Máximo	1275

Gráfico 3: Histograma para Quino.

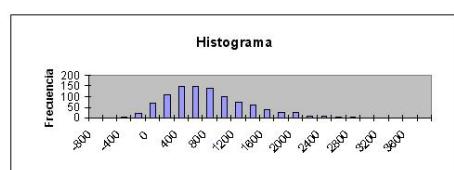


Fuente: Elaboración propia.

Tabla 7: Resultados simulación del VPN para Traiguén (millones de \$)

Media	692
Desviación estándar	604
Mínimo	-534
Máximo	3541

Gráfico 4: Histograma para Traiguén.



Fuente: Elaboración propia.

Para los dos proyectos los resultados ex post se encuentran dentro del histograma y la media de estas simulaciones es más similar al valor ex post que el valor determinístico (ver resumen en Tabla 9). Se confirman que la evaluación ex post entrega resultados que se asemejan más a los del modelo que las evaluaciones determinísticas previas. Con los resultados de estas calibraciones del modelo en los respectivos proyectos presentados, se considera validada la metodología no paramétrica y se valida la hipótesis de que se pueden estimar los riesgos de proyectos de embalses mediante modelos basados en VaR.

Tabla 9: Resumen de resultados de la calibración (Miles de \$ ó US\$):

Proyecto	VPN det.	VPN simulado	VPN ex post	VaR simulado
Victoria de Vilcún	U\$13.085	U\$12.210	U\$11.558	U\$8.000
Quino	-\$686.000	-\$860.000	-\$1.056.000	-\$1.471.000
Traiguén	\$923.000	\$692.000	\$430.000	-\$123.000

Fuente: Elaboración propia.

Una extensión del modelo: versión "jump-diffusion":

Se realizaron simulaciones adicionales (no presentadas en este documento) con un modelamiento de procesos incluyendo saltos por eventos extremos. En

el caso de proyectos de embalses para riego, se pueden considerar como eventos o caídas no consideradas en el Movimiento Browniano Geométrico, situaciones extremas de eventos climáticos (los embalses garantizan seguridad sólo en 85 de cada 100 años) como sequías prolongadas que afecten la disponibilidad de agua o inundaciones y rebalses que afecten las cosechas, también se podría tratar de cataclismos que dañen el embalse o las vías de transporte y comercialización de los cultivos. El proceso estocástico del logaritmo de los ingresos S es en este caso (Villamil, 2006):

$$d\ln(S) = (\mu - \sigma^2/2)dt + \sigma dW + \ln(J)dq \quad (14)$$

Donde dq toma valores 1 ó 0 con probabilidades λdt y $(1-\lambda)dt$. (proceso de Poisson). Esta es una versión "jump-diffusion" de Itô. Los resultados obtenidos con este modelamiento fueron aún mejores que los del modelo paramétrico presentados en la tabla 9.

7. Conclusiones

7.1. Objetivos, hipótesis y conclusiones finales

En relación a los objetivos iniciales de la investigación: Desarrollar una metodología de medición de rentabilidades y riesgos, basada en VaR, que permita realizar comparaciones entre proyectos del sector riego, y aplicar dicha metodología a proyectos de embalses, estos objetivos se cumplieron satisfactoriamente. La concreción de los objetivos anteriores permitió validar la hipótesis: se pueden estimar los riesgos de proyectos de embalses mediante modelos basados en VaR, en particular aquellos no paramétricos, que vía simulación de Monte Carlo permiten incluir además las opciones.

Respecto a la aplicabilidad del análisis de riesgo a proyectos de riego, las evaluaciones presentadas demuestran que la metodología propuesta con análisis de sensibilidad, el análisis probabilístico como marco teórico y la simulación como método práctico, es perfectamente aplicable, se comprobó además que parte importante de la información necesaria para la aplicación de la metodología está disponible y es pública (en particular los precios de productos agrícolas) y que en consecuencia es factible aplicar un enfoque no paramétrico de Value at Risk.

La metodología se calibró con los proyectos Victoria de Vilcún, Quino y Traiguén, obteniéndose muy buenos resultados predictivos, en el sentido de que la evaluación ex post entrega resultados que se asemejan más a los del modelo que las evaluaciones determinísticas previas. Con los resultados

de estas calibraciones del modelo, se considera validada la metodología no paramétrica desarrollada la hipótesis de la investigación.

Se puede aseverar que los resultados obtenidos muestran que las herramientas de análisis de riesgo aportan valor a la evaluación de proyectos: tanto en la aplicación a los proyectos Ligua y Petorca, como en los proyectos Quino y Traiguén, se pudo determinar la existencia de probabilidades de VPN negativo distintas de cero, lo cual constituye información valiosa para la toma de decisiones⁴.

La metodología y el modelo programado permiten además de cuantificar el Value at Risk (VaR) del proyecto, realizar una aplicación aproximada de opciones reales para valorizar la flexibilidad de abandonar los cultivos si los escenarios son demasiado desfavorables. Los resultados anteriores, permitirán una mejor priorización de proyectos. En efecto, si se generaliza la estimación de este tipo de métricas, será posible jerarquizar con más criterios: para proyectos de VPN similares, será posible discriminar según probabilidad de perder, VaR, intervalo de confianza, o una combinación de esos criterios. Estos nuevos criterios (actualmente las decisiones se toman basadas sólo en VPN) permitirán mejorar además el proceso de toma de decisiones de concesionar o no una obra de embalse.

Se comprobó que la medida de riesgo es mayor con VaR paramétrico. Este sesgo tiene que ver con la no valoración de las opciones reales que sí se consideran en el modelo no paramétrico. Dado que la aplicación paramétrica tomó como base los valores esperados del VPN de modelo no paramétrico, se puede afirmar que en el análisis comparativo las diferencias radican en la medición de riesgos (y no en la medición de rentabilidades).

Por último, se desarrolló un modelo de simulación (no paramétrico) que incluyó saltos de Poisson para capturar los efectos de eventos climáticos catastróficos. De acuerdo a los que se esperaba, los valores del VPN esperado y del VaR disminuyeron. Cabe señalar que en los dos proyectos analizados el modelo no paramétrico con eventos catastróficos (denominado modelo extendido) entregó resultados más cercanos al VPN ex post.

7.2. Nuevas líneas de investigación

Respecto a los pasos futuros, se puede mejorar un análisis que sólo se pudo desarrollar parcialmente en este estudio: la aplicación de opciones reales. Este último enfoque no fue posible de aplicarlo en forma rigurosa por falta de información respecto a la flexibilidad operacional que pudiera ser más interesante para proyectos de riego: la de cambiar la estructura de cultivos dependiendo de la evolución de los precios de los productos. No se contaba con informa-

⁴Este resultado también se verificó en el proyecto Ancoa (Contreras, 2008), que no se incluyó en este documento.

ción como para hacer esa valoración. En sustitución, se realizó un ejercicio valorando la opción de sustituir por cultivos que entregaran $VPN = 0$.

En lo que respecta al modelamiento de precios, dada la existencia de tendencias en los retornos de los precios, se podría explorar la alternativa de la aplicación de un proceso de Metcalf & Hassett (en lugar del movimiento browniano geométrico) que considera también tendencia y además reversión a la media. Adicionalmente, se puede modelar la evolución de los parámetros μ y σ a lo largo del tiempo (tratados entonces como μ_t y σ_t), con coeficientes de reversión a la media para dichos parámetros, y modelando procesos para estas variables en los primeros años del horizonte de evaluación. Por lo tanto una línea completa de investigación puede pasar por el análisis de procesos tipo Metcalf y Hassett, o modelos como los de Schwartz y Moon (2001).

Por último, una línea complementaria de investigación, consiste en determinar la estructura óptima de cultivos con modelos de portafolio (Markowitz, 1952). En esta investigación se trabajó con las actuales estructuras de cultivos (obtenidas de estudios de consultores que levantan dicha información con encuestas y trabajo de terreno), la alternativa es estimar una estructura óptima según rentabilidad y riesgo, a la cual los agricultores podrían tender en el futuro si se mantuviesen los niveles de riesgo y rentabilidad (o la que tendrían con los niveles de riesgo y rentabilidad estimados según los procesos).

Referencias

- [1] Alessi, G. (2005) Value at Risk (VaR) in Real Options analysis. Università degli Studi - L'Aquila. Facoltà di Economia. Dip. di Sistemi ed Istituzioni per l'Economia.
- [2] Ayala y Cabrera (2007) Estudio de Factibilidad Habilitación La Ligua -Petorca, Informe Técnico.
- [3] Best P. (1999). Implementing Value at Risk. John Wiley & Sons Ltd.
- [4] Blanck Leland y Tarquin Anthony.(1998) Ingeniería Económica. 3a. ed. McGraw Hill.
- [5] Carmona, S. (2007) Cuantificación del efecto de la incertidumbre de precios de cultivos en la evaluación económica de un proyecto de riego. Memoria de título de Ingeniero Civil Industrial. Universidad de Chile.
- [6] Cartes, Contreras y Cruz (2005) La Tasa Social de Descuento en Chile. Documentos de Trabajo. Serie de Gestión. N°77. Departamento de Ing. Industrial. Universidad de Chile.

- [7] Contreras E, y Fernández V. (2003). Una nueva metodología para la evaluación de proyectos de inversión: Las Opciones Reales. Revista de Ingeniería de Sistemas. Volumen XVII, Número 1, Julio de 2003.
- [8] Contreras E (2008) Aplicación de metodologías basadas en VaR para el análisis de riesgo en proyectos de embalses. Tesis Doctoral. Universidad Autónoma de Madrid.
- [9] Dowd K. (1999) Beyond Value at Risk - The new science of risk management. John Wiley & Sons Ltd. Reprinted January 1999.
- [10] Embrechts, P. H̄oing, A. Juri A. (2002) Using Copulae to bound the Value-at-Risk for functions of dependent risks. Department of Mathematics ETHZ, CH-8092 Zurich, Switzerland.
- [11] Engel y Valdés (2002) Prediciendo el precio del cobre: ¿Más allá del camino aleatorio?.En Dilemas y Debates en torno al cobre. Dolmen.(Patricio Meller editor)
- [12] Ergashev, B. (2002) A note on a generalized Black-Sholes formula. Department of Economics. Elliot Hall Campus.
- [13] Feria, J.M. (2005) El riesgo de Mercado: su medición y control. España. Delta Publicaciones.
- [14] Fernández V. (2007), Procesos no Estacionarios, Apuntes del curso IN72W, Departamento de Ingeniería Industrial de la Universidad de Chile.
- [15] Figueroa G, Muñoz V et al (1993) .“ Regadío Canal Victoria - Traiguén - Lautaro” Programa de Adiestramiento en preparación y evaluación de proyectos, Santiago - Chile. Instituto de Economía. Pontificia Universidad Católica.
- [16] Glasserman, Paul, Heidelberger Philip and Shahabuddi Perwez. (2000) Efficient Monte Carlo methods for value-at-risk.
- [17] Jorion P. (2004) Value at Risk: el nuevo paradigma para el control de riesgos con derivados. Editorial Limusa. México 2004.
- [18] Lamothe Prosper y Otero Jorge (2003) Garantías de Valor Residual en la Industria Aeronáutica:Un Enfoque de Valoración Basado en Opciones Reales. Universidad Autónoma de Madrid.
- [19] Lamothe P. y Pérez Somalo (2004). Opciones Financieras y Productos Estructurados España. Mc Graw Hill. Segunda Edición.

- [20] Longerstaey J. & Zangar P, Market risk research group at JP Morgan, (1995) .^A transparent tool”, Market Risk, Vol. 8, N° 1.
- [21] Metcalf y Hasset, (1995). Investment Under alternative return assumptions: comparing random walk and mean reversion. National Bureau of Economic Research.
- [22] Markowitz, Harry. (1952). “Portfolio Selection”, Journal of Finance, 7: pp. 77 -91.
- [23] Schwartz E. y M. Moon, (2001), “Rational Pricing of Internet Companies Revisited.” Documento de trabajo, Anderson School of Business, University of California.
- [24] Stein J., Usher S., LaGattuta D. and Youngen J, (2001), “A Comparable Approach to Measuring CashFlow-At-Risk for Non-Financial Firms” . National Economic Research Associates, Journal of Applied Corporate Finance. Vol 13. N° 4, 8-17.
- [25] Vilariño A. (2001) Turbulencias financieras y riesgos de mercado. Financial Times - Prentice Hall.
- [26] Villalpando L. (2007) Simulación de precios del petroleo brent. Disponible en <<http://detodoynada.nireblog.com/blogs1/detodoynada/files/trabajo-procesos.pdf>>.
- [27] Villamil J. (2006) Modelos de valoración de opciones europeas en tiempo continuo. Cuadernos de Economía. ISSN 0121-4772. v.25 n.44 Bogotá.

SISTEMA DE MONITOREO DE MERCADO: APLICACIÓN AL SECTOR ELÉCTRICO

RODRIGO PALMA BEHNKE*

SHARO ESCOBAR*

TOMÁS REID*

Resumen

En un mercado específico, las variables de interés cambian dinámicamente. Si los actores que participan no se informan de los cambios oportunamente, las decisiones que toman pueden traducirse en una asignación ineficiente de los recursos. A modo de ejemplo, la introducción de competencia en el sector eléctrico ha motivado el desarrollo de herramientas para supervisar el mercado, de manera de detectar conductas anticompetitivas. Asimismo, los agentes de un mercado hacen uso de distintas fuentes de información e indicadores que les permiten desarrollar sus estrategias de negocios. En este sentido, el presente trabajo propone un esquema general para elaborar un sistema de monitoreo de mercado, evaluando su desempeño en una aplicación realizada al sector eléctrico.

Los pasos esenciales para construir un sistema de monitoreo de mercado son: identificar los requerimientos del negocio, modelar los datos adecuadamente, diseñar una arquitectura para editar y mostrar los datos, implementar el sistema en un lenguaje de programación adecuado y distribuir la información a los actores. Para lograr informar oportunamente sobre la situación del mercado, los resultados calculados deben mostrarse en un ciclo de monitoreo de corta duración enfocado en los elementos de interés. Para darle validez a los resultados, la información debe certificar su calidad mediante técnicas de Data Quality.

Palabras Clave: Monitoreo de Mercado, Mercado Eléctrico, Data Quality.

*Departamento de Ingeniería Eléctrica, Universidad de Chile

1. Introducción

1.1. Motivación y Objetivo

En los sectores productivos las variables de interés van cambiando dinámicamente, lo que hace necesario que los actores participantes se mantengan informados de lo que ocurre. Dicho de otra forma, una *visualización del mercado* es esencial como apoyo al proceso de toma de decisiones. Los *newsletters* que ofrecen empresas de información, pese a entregar un resumen de las noticias diarias de un sector productivo, filtradas con procesos de selección especializados, no permiten responder cabalmente a las necesidades de diagnóstico específicas de un agente de mercado. A modo de ejemplo, a través de estos servicios no es posible realizar de manera simple el cálculo temporal de un cierto indicador.

Sin embargo, en algunos sectores se han desarrollado plataformas de visualización específicas. Este es el caso de entidades bancarias y centros de inversión, donde pantallas muestran la evolución de las principales acciones, los indicadores macroeconómicos y aquellos representativos del estado de evolución de una bolsa de comercio, etc.

Esta idea desarrollada inicialmente en el mundo financiero es susceptible de ser replicada en otros sectores productivos, creando indicadores específicos que sean de utilidad para los usuarios y que les permitan tomar decisiones oportunas y con una adecuada base de información. Desde una perspectiva de evaluación social, este tipo de mecanismos contribuye a una mejor asignación de los recursos de un sistema.

En caso del sector eléctrico, se utiliza un volumen importante de información, tanto de la operación como de los resultados económicos asociados, lo que motiva la creación de indicadores que resuman la información así como de plataformas de visualización adecuadas, de modo de crear un soporte a las decisiones operativas y de inversión de los actores participantes.

En este contexto, el presente trabajo propone un esquema general para el diseño e implementación de un sistema de monitoreo de mercado, evaluando su desempeño en una aplicación realizada al sector eléctrico.

En el siguiente punto se presenta una visión general del sector eléctrico. En la sección dos se presentan las bases de un sistema de Monitoreo de Mercado. En la sección tres se presenta la propuesta para la construcción de un sistema de monitoreo de mercado. En la sección cuatro se presentan aspectos de la implementación del sistema. Por último, en la sección cinco se entregan las conclusiones del trabajo.

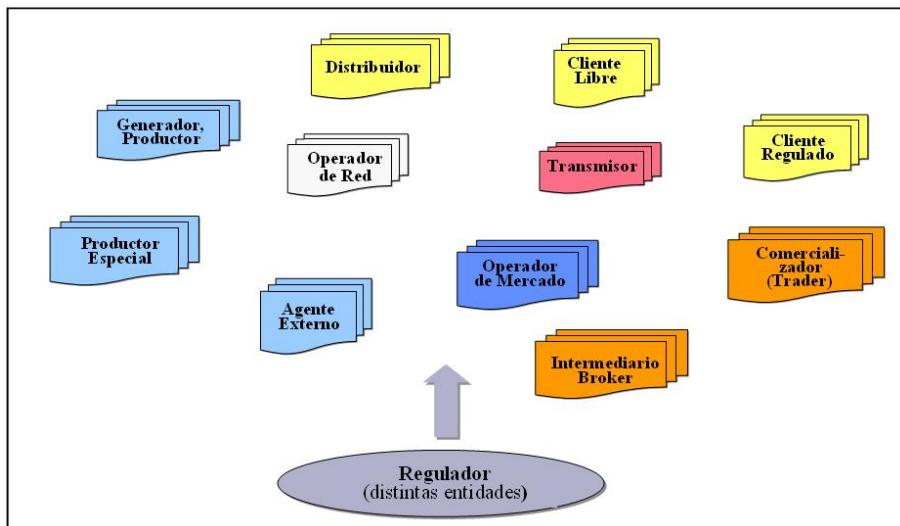


Figura 1. Actores del mercado eléctrico.

1.2. El Mercado Eléctrico

Un sistema eléctrico de potencia SEP es un conjunto de instalaciones que permiten generar, transportar y distribuir la energía eléctrica en condiciones adecuadas de tensión, frecuencia y disponibilidad. Este SEP permite el intercambio de productos eléctricos de forma no discriminatoria. Estos productos son: energía, potencia y servicios complementarios. Los países han desarrollado distintos modelos de mercados competitivos que permiten el intercambio de algunos de dichos productos, siendo Chile el primer país que introdujo competencia en el sector.

Participan en el sector eléctrico distintos actores, los cuales deben examinar el mercado de modo de buscar oportunidades de negocio (figura 1). Los generadores, productores y agentes externos como conexiones internacionales forman el grupo de los productores. Las empresas distribuidoras, clientes libres y clientes regulados forman el grupo de consumidores. Un actor especial es el transmisor, que permite el intercambio de los productos entre productores y consumidores. Como agentes del negocio están los broker y trader. El encargado de que todo el mercado funcione en condiciones adecuadas es el operador de red, cuyo campo de acción queda limitado por las reglas impuestas por los reguladores.

2. Sistema de monitoreo de mercado - SMM

Como se observa en la figura 2, un sistema de monitoreo de mercado se define como una entidad destinada a vigilar variables de riesgo y elaborar

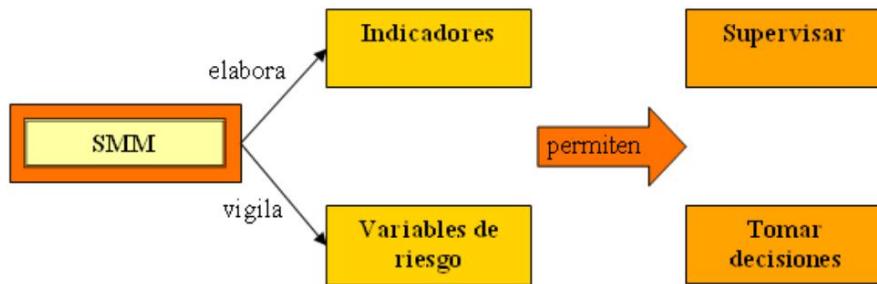


Figura 2. Definición de sistema de monitoreo de mercado.

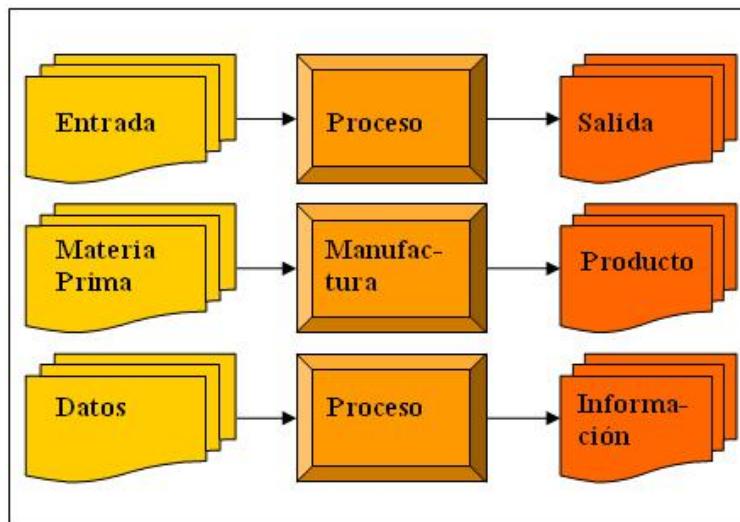


Figura 3. Analogía entre procesamiento de materias primas y procesamiento de datos [2].

índices de desempeño que permitan visualizar conductas anticompetitivas y oportunidades de inversión, verificando si se realiza una asignación eficiente de los recursos disponibles.

2.1. Bases de la calidad de información

En un sistema de monitoreo de mercado, la materia prima son los datos y los productos son la información (2). Esto se observa en la figura 3.

Para entender la información como un producto de calidad, debe ser (10): accesible, interpretable, contextualizada e intrínseca. Considerando las dimensiones anteriores, se puede definir una visión y principios para un sistema de información como lo realiza la U.S. Security and Exchange Comision en (4) y el Australian Biodiversity Information Services en (3).

En cuanto a la visión que debe tener un sistema de monitoreo de mercado, éste debe corresponder a una herramienta de soporte a la supervisión y toma de decisiones en el sector al cual esté enfocado, basada en datos confiables,

respaldados con fuentes adecuadas y con resultados claros, reproducibles y justificados.

Los principios que debe tener el SMM son:

Utilidad: Responder interrogantes de los usuarios y ser alimentado por ellos mismos (feedback).

Objetividad: Tener indicadores claros, exactos, autocontenidos (completos) e imparciales.

Integridad: Los datos deben ser consistentes respecto a sus atributos y se deben documentar las fuentes.

Transparencia: Los errores de consistencia deben estar disponibles mediante alarmas automáticas.

3. Propuesta. Pasos para la construcción de un SMM

Para tener un sistema de monitoreo de mercado debe construirse un almacén de datos sobre el cual se calculen indicadores los cuales se visualicen en una plataforma adecuada. Estos datos son integrados con más frecuencia por las empresas para obtener información exacta, a tiempo y útil (5). En la figura 4 se observan los pasos propuestos para la construcción de un sistema de monitoreo de mercado.

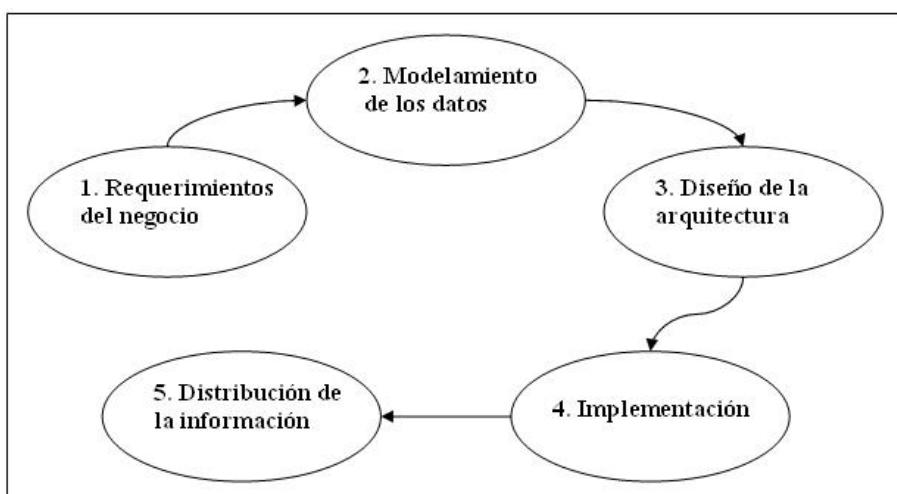


Figura 4. Pasos para la construcción de un sistema de monitoreo de mercado [5].

3.1. Requerimientos del negocio

Corresponde a identificar las preguntas del negocio desde los usuarios del sistema de monitoreo de mercado. Se consigue esta información mediante encuestas y jornadas de trabajo en grupo. Es necesario saber:

- Qué quieren ver.
- Qué tipo de trabajos realizan.
- Qué desafíos que enfrentan.
- Qué tipo de decisiones deben tomar.

Al poder ser muchos y distintos los usuarios de un sistema de monitoreo de mercado, la alternativa más viable para definir los requerimientos es la realización de una encuesta, ejemplificada en la figura 5. Los resultados permiten identificar las variables e indicadores de mayor interés de modo que el SMM cumpla con el principio de utilidad.

3.1.1. Variables de riesgo e indicadores en el caso eléctrico

Las variables de riesgo vigiladas por el proceso de monitoreo de mercado en el caso eléctrico y los indicadores elaborados deben cumplir con los siguientes criterios (1):

- Permitir la realización de un adecuado análisis a las empresas participantes y del mercado como un todo.

	Muy importante	Importante	Indiferente	Poco importante	Irrelevante
Variable 1	*				
Variable 2			*		
etc...					

Figura 5. Ejemplo de encuesta para identificar los requerimientos del negocio.

- Ser factibles de implementar desde el punto de vista de la adquisición de información o en base a información estimada con una adecuada precisión.
- Ser la cantidad estrictamente necesaria.

Considerando lo anterior, los indicadores propuestos se han organizado en función al tipo de análisis que realizan:

Situación de operación del sistema: En esta categoría se agrupan parámetros sobre la operación del sistema, con la finalidad de supervisar su evolución. Ejemplos de ellos son: demanda, reserva, congestiones, operación, despacho real v/s programado, cotas de embalses, mantenimiento de unidades, estadística de fallas, etc.

Situación de operación del mercado: Agrupa parámetros del mercado en sí, como precio spot, precio de combustibles, costos de operación de centrales térmicas, costo de oportunidad de embalses, etc.

Estructura del mercado: Muestran las características estructurales del sistema que pudiesen configurar la existencia de circunstancias para el ejercicio de poder de mercado. Ejemplos de ellos son la capacidad instalada y la concentración del mercado.

Desempeño y poder de mercado: Evalúan el desempeño y tratan de establecer la existencia de ejercicio de poder de mercado reflejando la existencia de comportamiento competitivo.

Una lista exhaustiva de los indicadores propuestos para cada grupo se encuentra en (1).

3.2. Modelamiento de los datos y aplicación al sector eléctrico

Para construir los indicadores propuestos en la sección anterior, es necesario que los datos de entrada se guarden en una estructura que se adapte al sistema de monitoreo de mercado. Para el caso eléctrico se propone el siguiente diseño conceptual y lógico de los datos de entrada.

DISEÑO CONCEPTUAL: Se requiere identificar qué necesita de los datos la BD para satisfacer los requerimientos del negocio, independiente del sistema utilizado para manejar las bases de datos. Como resultado de este proceso se logra un diagrama de relaciones como en una clase orientada al objeto.

En el sector eléctrico el grupo propone dividir los datos de entrada en 3 bases de datos: red eléctrica, mercado eléctrico y red hidráulica, que en el caso de países dependientes del agua para la generación como Chile, esta última resulta de suma importancia. El diagrama de relaciones obtenido para la red eléctrica y el mercado se observa en la figura 6. Por ejemplo, un generador es un tipo de inyección, que a la vez es un elemento de 1 polo, que a la vez es un elemento de la red, que a la vez es un elemento del sistema. Así, un generador hereda todos los atributos de sus clases superiores.

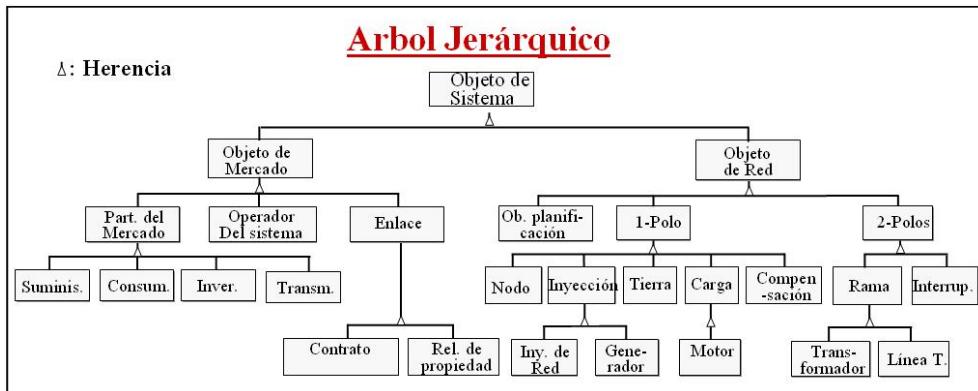


Figura 6. Diagrama de relaciones de los elementos del SMM en el caso eléctrico.

En la figura 7 se muestran los atributos de cada objeto perteneciente a la base de datos de la red. Por ejemplo, una relación de propiedad tiene los atributos propios y heredados de las clases superiores que se observan en el cuadro 1.

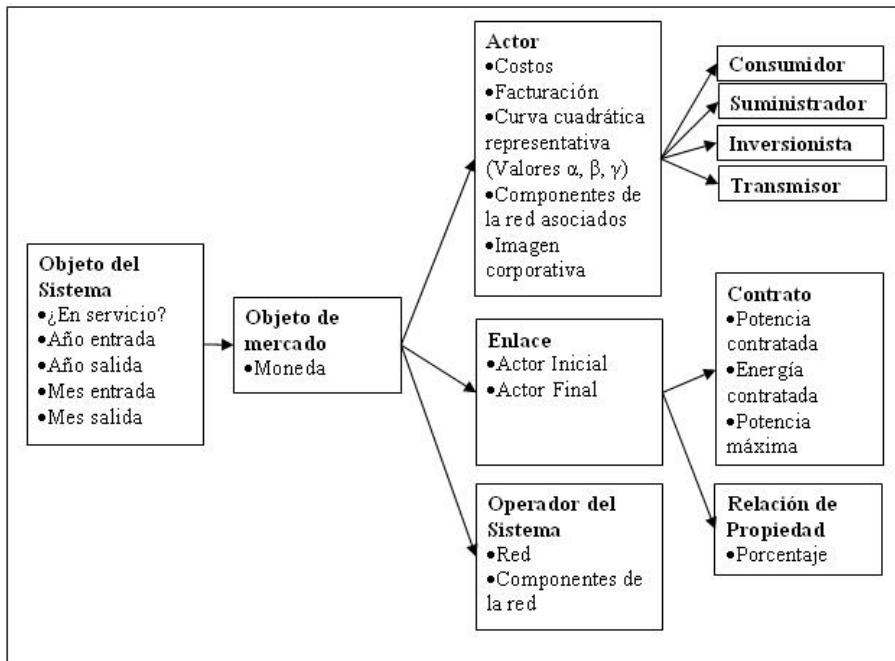


Figura 7. Ejemplo de los campos necesarios para describir cada elemento del editor de mercado, basado en la herencia de los campos de las clases superiores.

DISEÑO LÓGICO: Se lleva lo anterior al formato en que serán almacenados. Para el almacenamiento y la carga de esta base de datos orientada al objeto, se hace uso de archivos fuentes en formato ASCII y bases de datos relacionales. La información contenida en la base de datos puede ser modificada

- ¿En servicio?	- Moneda
- Año entrada	- Actor inicial
- Año salida	- Actor final
- Mes entrada	- Porcentaje
- Mes salida	

Cuadro 1: Atributos del objeto *relación de propiedad*.

en línea, a través de los editores de red y de mercado.

3.3. Diseño de la arquitectura

Una arquitectura de software consiste en un conjunto de patrones y abstracciones coherentes que proporcionan el marco de referencia necesario para la guiar la construcción del software para un sistema de información (8).

El modelamiento de los datos orientado al objeto descrito en la sección anterior, ha sido implementado como un sistema distribuido basado en una estructura de tipo cliente-servidor. Una estructura cliente - servidor supone que algunas funciones se realizan mejor en procesos locales y otras es mejor realizarlas de manera centralizada (7). La figura 8 muestra la arquitectura del paquete de programación desarrollado. Las flechas grises representan servicios requeridos por parte de los clientes a los respectivos servidores, en tanto que las de color negro simbolizan un mero intercambio de datos. Las bases de datos BDR, BDM y BDH constituyen la parte medular del sistema, siendo sus servicios solicitados por todas las componentes del sistema. De esta forma, se permite un acceso controlado a la información de los objetos del sistema. La información contenida en la base de datos puede ser modificada en línea, a través de los editores de red, de mercado e hidráulico. Estos editores, junto con un sistema de información geográfica, disponen de interfaces gráficas que permiten una comunicación flexible hombre-máquina. La biblioteca manejadora de eventos posibilita, tanto el uso de los dispositivos de entrada y salida como la capacidad de interactuar con la base de datos a través de protocolos y servicios internet (6).

4. Implementación del SMM

4.1. Estructura general

De acuerdo a la propuesta general de diseño presentada en la sección anterior, es factible comenzar la etapa de implementación, correspondiente al

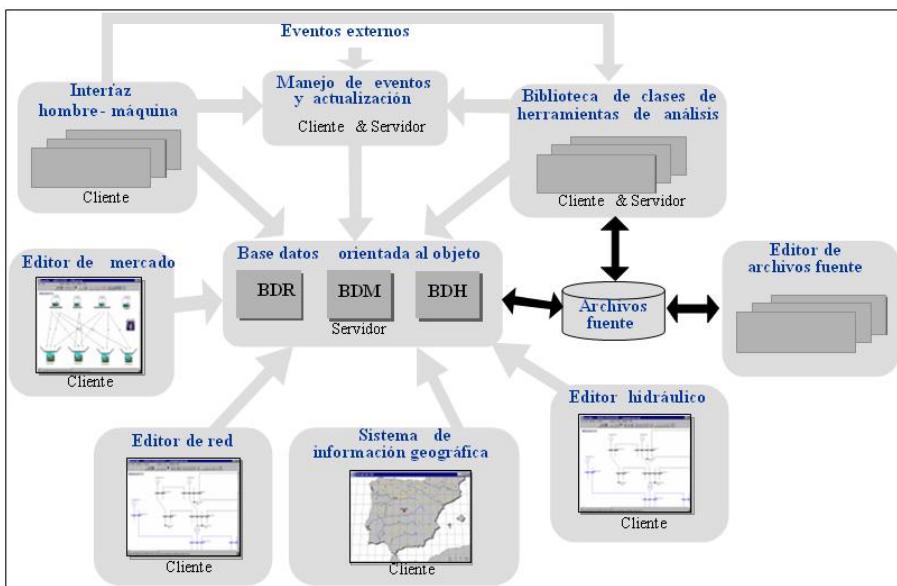


Figura 8. Arquitectura del sistema de monitoreo de mercado.

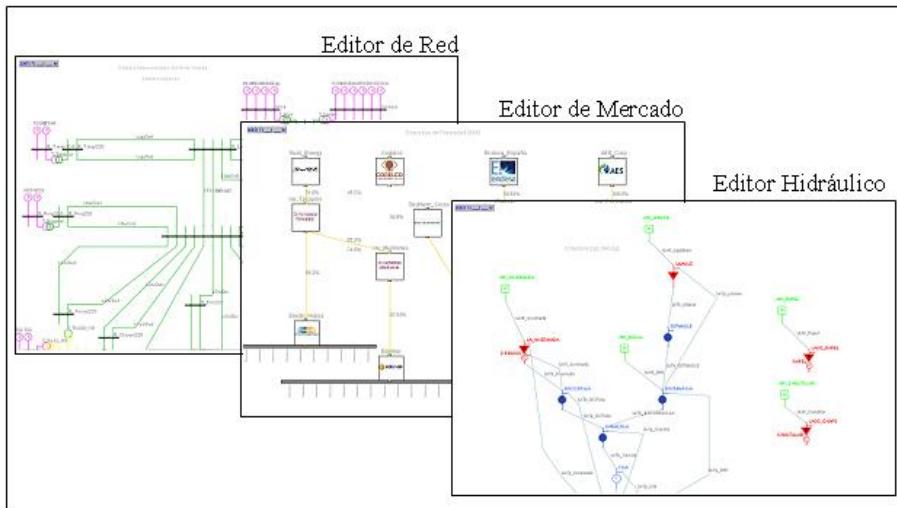


Figura 9. Editores del programa.

montaje sobre una plataforma preexistente. La arquitectura planteada se programa en el lenguaje Java ya que está orientado al objeto lo que hace fácil modelar los datos según lo planteado. En la figura 9 se observan los editores de red, de mercado e hidráulico desarrollados. Estos editores sirven para ingresar la información a las bases de datos de manera gráfica. También se puede ingresar información a través de algún programa que pueda operar con bases de datos como MS ACCESS, lo que permite el manejo de los grandes volúmenes de información de las fuentes, como las identificadas en la figura 10 para el

caso eléctrico.

En esta etapa es importante incluir sistemas de validación que evalúen la rigurosidad y completitud de los datos (3). Así, medidas de calidad de datos implementadas en el sistema de monitoreo de mercado eléctrico son:

- Rangos de tiempo en que son válidos los datos, basado en la idea de que cada dato tiene un ciclo de vida o vigencia.
- Alarmas automáticas frente a inconsistencias de los datos (ej: la suma de participación de los dieños de una empresa determinada resulta mayor a 100 %).
- Registro de las fuentes en una base de datos y posibilidad de consultarlas.

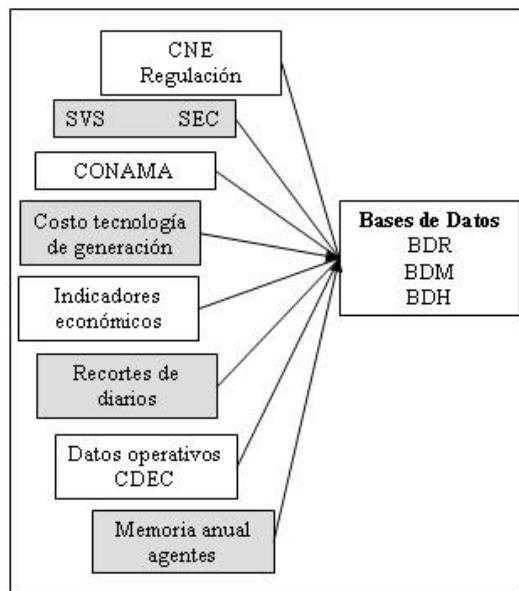


Figura 10. Fuentes de información de las bases de datos en el sistema de monitoreo del mercado eléctrico.

4.2. Distribución de la información

Para la distribución de la información se utilizan herramientas de visualización y reportes adecuados. Existen un sinnúmero de alternativas para mostrar información: tablas, números, gráficos (de columnas, de líneas, circulares, de barra, etc.), colores, agujas, flechas, animaciones, películas, contornos de colores, entre otros, algunos de los cuales se adaptan de mejor forma a un

mercado eléctrico (9). En la sección siguiente se plantea una forma de distribuir la información a través de un ciclo de monitoreo.

Visualizar todas las variables e indicadores en todos los puntos del sistema disponible no resultaría práctico. Luego la herramienta de monitoreo de mercado debe tener una aplicación que recorra el mercado con una ventana de tiempo móvil, sobre elementos de interés.

En el caso mostrado en la figura 11, se elige el tiempo de inicio y fin del monitoreo junto con el tamaño de la ventana. En la misma ventana, se elige qué variables se quiere monitorear, el tipo de visualización y los elementos de interés. Al comenzar el ciclo, el tiempo se actualiza al año y mes de inicio para luego recorrer en forma secuencial el editor de red donde se muestra la evolución de los precios en 3 barras, el editor de mercado donde se ve la repartición de la energía generada entre 5 empresas y el editor hidráulico donde se muestra con un gráfico de aguja el nivel de un embalse. Una vez recorrido los 3 editores, se actualiza el tiempo del sistema, redibujando los gráficos correspondientes para mantener la ventana de tiempo seleccionada.

Esta propuesta puede ser aplicada en términos prácticos a los sistemas nacionales como es el caso del sistema interconectado central SIC (11)

5. Conclusiones

En este trabajo se propone una forma de elaborar sistemas de monitoreo de mercado, basado en técnicas de extracción de conocimiento de bases de datos. La aplicación general al caso del mercado eléctrico busca validar los conceptos propuestos.

Se propone una visión y un listado de principios que debe cumplir un SMM basado en las dimensiones de la teoría de calidad de datos, de modo de certificar la información. Los principios son utilidad, objetividad, integridad y transparencia.

El primer paso de la construcción de un SMM pasa por identificar los requerimientos del negocio. En el caso eléctrico, se han identificado variables divididas en 4 grupos: situación de operación del sistema; situación de operación del mercado; estructura del mercado; y desempeño y poder de mercado.

El siguiente paso es modelar los datos. Para el caso eléctrico se dividen los datos en 3 grupos: base de datos de la red, base de datos del mercado y base de datos hidráulico, construyendo un árbol jerárquico con atributos heredados que permitan la organización correcta y la posterior implementación. Luego se propone una arquitectura cliente - servidor que permite que aplicaciones interactúen con las bases de datos simplificando la programación.

La implementación se realiza en el lenguaje Java ya que es un lenguaje

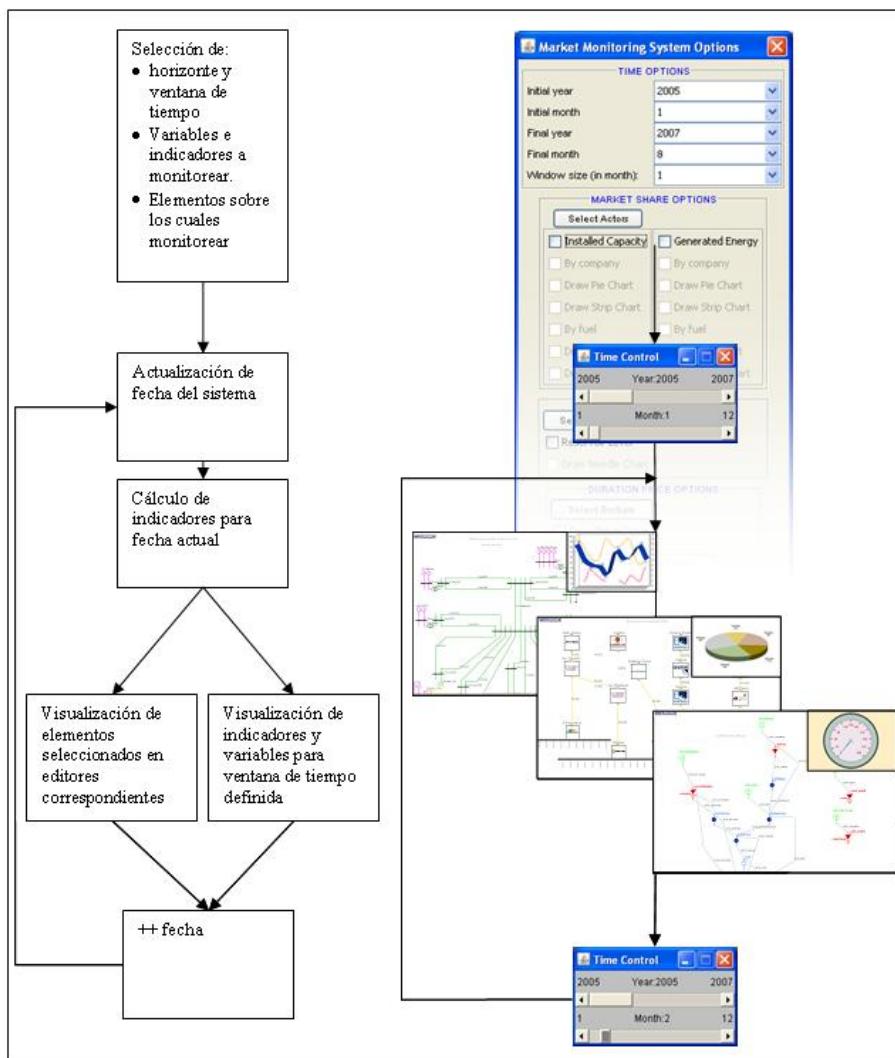


Figura 11. Diagrama de flujo del ciclo de monitoreo.

orientado al objeto y simplifica la programación con el modelamiento de los datos propuestos. Para la distribución de la información se utilizan los editores de las bases de datos interactuando con formas de visualización gráficas.

Por último, se plantea un ciclo de monitoreo que corre sobre el sistema de monitoreo de mercado, visualizando en el horizonte de tiempo seleccionado y sobre los elementos deseados, las variables e indicadores de interés, logrando que el usuario en minutos se informe del estado del mercado.

Agradecimientos: Al Instituto Milenio: *Sistemas Complejos de Ingeniería* por el apoyo brindado en la elaboración y financiamiento de este trabajo.

Referencias

- [1] Escobar, S. *Sistema de monitoreo de mercado para el análisis de riesgo en el segmento de generación eléctrica en sistemas hidrotérmicos*. Tesis magíster en ciencias mención ingeniería eléctrica. Borrador, 2007.
- [2] Wang, R., Storey, V., Firth, C. *A framework for analysis of data quality research*. IEEE transactions on knowledge and data engineering, Vol , No. 4. August 1995.
- [3] Chapman, A. D. 2005. *Principles of data quality, version 1.0*. Report for the global biodiversity information facility, Copenhagen.
- [4] SEC 2002. *Final data quality assurance guidelines*. United States Securities and Exchange Commission. www.sec.gov/about/dataqualityguide.htm Consultado el 27 de Agosto de 2007.
- [5] Sen, A., Sinha, A. *Toward developing data warehousing process standards: An ontology-based review of existing methodologies*. IEEE transactions on man and cybernetics. Vol.37, NO. 1. January 2007.
- [6] Handschin, E., Müller, L., Nikodem, L., Palma, R. *Modelo Orientado al Objeto para la simulación y gestión de Mercados Eléctricos Competitivos*. IEEE-Andescon 99, Isla Margarita Venezuela, 1999.
- [7] Davey, B., Tatnall, A. *Tools for client - server computing*. International conference on software engineering: education and practice, 1996.
- [8] Wikipedia. *Arquitectura software*. http://es.wikipedia.org/wiki/Arquitectura_software Consultado el 24 de Septiembre de 2007.
- [9] Laufenberg, M.J. *Visualization aproaches integrating real time market data*. Power Systems Conference and Exposition, Octubre 2004. IEEE PES.
- [10] Mielke, M. *Information Quality principles and foundations*. Taller de calidad de datos, la base para una mejor toma de decisiones. Santiago, Chile. Septiembre 2007.
- [11] T. Reid M., R. Palma-Behnke, S. Escobar *Sistema de Monitoreo de Mercado Aplicación al Sector Eléctrico*. Anales del Instituto de Ingenieros, diciembre, 2007.

PROGRAMACIÓN DE HORARIOS DE CLASES Y ASIGNACIÓN DE SALAS PARA LA FACULTAD DE INGENIERÍA DE LA UNIVERSIDAD DIEGO PORTALES MEDIANTE UN ENFOQUE DE PROGRAMACIÓN ENTERA

RODRIGO HERNÁNDEZ*

JAIME MIRANDA P.**

PABLO A. REY***

Resumen

Un aspecto importante en la gestión académica de las universidades es la generación de horarios y la asignación de salas de clase para los distintos cursos que realizan. En este artículo se presenta un modelo de programación entera el cual decide simultáneamente los horarios de los cursos y la asignación de salas. Las variables utilizadas están asociadas a la definición del horario del curso para una semana por medio de un patrón horario. Una particularidad del modelo es que tanto las condiciones sobre capacidad y tipo de salas de clase, así como las combinaciones de bloques horarios para un curso, son manejadas implícitamente mediante las variables de decisión. Se reportan los resultados da la comparación de la programación obtenida con el modelo propuesto y la programación que efectivamente se utilizó. El modelo propuesto entrega de manera rápida y eficiente los horarios y asignaciones de sala de clase satisfaciendo todos los requerimientos obligatorios y condiciones deseables para la Facultad en un tiempo menor a los 5 minutos.

Palabras Clave: *Timetabling, class scheduling, programación entera.*

*Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.

**Departamento de Control de Gestión y Sistemas de Información, Facultad de Economía y Negocios, Universidad de Chile, Santiago, Chile.

***Escuela de Ingeniería Industrial, Facultad de Ingeniería, Universidad Diego Portales, Santiago, Chile.

1. Introducción

Las instituciones educacionales enfrentan cada semestre el problema de la programación de horarios y asignación de salas de clase de los cursos que imparten. Desde la perspectiva de la Investigación de Operaciones, este tipo de problemas se enmarcan dentro del área conocida como *Timetabling* o programación horaria. Los problemas de esta área consisten en la asignación de ciertos eventos a distintos bloques horarios respetando una serie de requerimientos y condiciones. Dentro de estos problemas existe una rama específica, llamada *Class Scheduling*, que estudia problemas relacionados con la programación horaria para entidades educativas. Dentro de este contexto, existen tres tipos de problemas [45]:

- Programación de horarios de evaluaciones y exámenes (*Examination Timetabling*).
- Programación de horarios de clases para colegios (*School Course Timetabling*).
- Programación de horarios de clases para instituciones de educación superior o universidades (*University Course Timetabling*).

Como ya hemos señalado, la Facultad de Ingeniería debe resolver el problema de generar la programación horaria de sus cursos y la asignación de salas de clase, o sea, resuelve un problema del tipo *University Course Timetabling*.

Esta programación debe satisfacer una serie de requerimientos impuestos por políticas de la Facultad. Actualmente, el proceso de programación se realiza de forma manual demorando aproximadamente un mes en promedio. Cabe destacar que, la programación obtenida de esta manera no está libre de errores, detectándose en algunos casos ciertas inefficiencias e incumplimientos de los requerimientos básicos para el correcto funcionamiento de la institución educativa.

En consecuencia, una buena programación horaria genera una serie de beneficios para los principales actores que conviven en esta institución. Entre estos es posible mencionar por ejemplo: eliminar topes de horarios entre cursos del mismo semestre, respetar la disponibilidad de horarios de los profesores, respetar la capacidad de las salas de clase e incorporar condiciones deseables, como por ejemplo: favorecer las clases en bloque horarios específicos o minimizar la utilización de salas especiales.

En este trabajo, se propone un modelo de programación entera para la generación de la programación horaria y asignación de salas. Este modelo

incorpora, por un lado, todos los requerimientos y condiciones deseables y, por otro, los objetivos perseguidos por la Facultad.

Siguiendo esta línea, la estructura de este artículo es la siguiente: la sección 2 presenta los antecedentes relevantes para esta investigación. Primero, se expone un análisis bibliográfico en relación al problema de programación horaria presentando los principales enfoques de solución. Luego, se realiza una descripción del problema presentando sus principales características y requerimientos. La sección 3 detalla el modelo de programación entera. La sección 4 analiza los resultados computacionales de la implementación del modelo y se compara, en términos de indicadores de desempeño, con la actual forma de operar para el semestre Otoño 2007. Finalmente, la sección 5 presenta las conclusiones de este estudio.

2. Antecedentes

2.1. Análisis Bibliográfico

Los problemas de generación de horarios y asignación de recursos en instituciones educativas han sido ampliamente estudiados en la literatura [5, 44, 45]. Estos problemas pueden ser clasificados de acuerdo al tipo de institución educativa (colegios o universidades) y por el tipo de eventos a programar, clases o evaluaciones. Calendarizar clases en colegios o universidades son problemas bien diferentes en la práctica. Usualmente, en los colegios los alumnos que pertenecen a un determinado curso toman en bloque las mismas asignaturas, pues se desean horarios compactos de clases¹. En el caso de las universidades en cambio, debe existir cierta flexibilidad en los horarios y en la selección de los cursos que toma cada estudiante. La programación de clases y evaluaciones difieren, principalmente, en los aspectos siguientes [13]:

1. Las evaluaciones deben ser programadas después de la inscripción de los alumnos, mientras que los cursos son usualmente programados con anterioridad a la inscripción de los alumnos.

2. Las restricciones asociadas al uso de salas de clase pueden ser diferentes.

Las clases de un curso deben ser programadas, por lo general, en una misma sala de clase. Para las evaluaciones los requerimientos pueden ser diferentes, ya que, en algunos casos, se programan evaluaciones de diferentes cursos compartiendo la misma sala de clase y, en otros casos, se requiere programar las evaluaciones de un mismo curso en varias salas de clase debido a la cantidad de alumnos inscritos en dicho curso.

¹Horarios con clases seguidas o sin ventanas temporales entre cursos sucesivos.

A continuación se realiza una revisión bibliográfica de los trabajos publicados en esta área. La literatura presenta numerosas variaciones del problema de programación de horarios de acuerdo a los requerimientos específicos de cada institución. En particular, más allá de las restricciones de topes en la programación de horarios de profesores y uso de salas de clase, no hay otras condiciones que aparezcan en todos los estudios presentados en la literatura [33]. Sin embargo, existe una gran variedad de enfoques de solución para la generación de horarios y asignación de salas de clase. Mientras algunos trabajos se concentran en los aspectos prácticos y el desarrollo de sistemas, otros lo hacen en el modelamiento y en metodologías de solución.

Según las características de la institución, la programación de los horarios se realiza con antelación a la inscripción de los alumnos, o luego de ésta. Entre los trabajos que analizan el problema después que los alumnos han escogido sus cursos podemos mencionar el trabajo reciente de Boland *et al.* [9] en donde no se considera la capacidad de las salas de clase, sino que se definen un número predeterminado de secciones en cada curso y un número máximo de estudiantes por sección. Adicionalmente, en otros trabajos consideran la programación de los horarios de los cursos después de la inscripción de los alumnos [4, 18, 27].

En nuestro caso particular, la programación se debe realizar con anterioridad a la inscripción de los alumnos. En este caso, se deben imponer condiciones sobre los topes de horarios de los cursos que compartirán alumnos de acuerdo a estimaciones de demanda o al plan de estudios. Como veremos en la próxima sección, en nuestro problema, las condiciones impuestas corresponden a evitar topes horarios entre los cursos del mismo semestre, de acuerdo a los planes de estudio de las carreras que imparte la Facultad.

Las condiciones o requerimientos del problema pueden ser clasificados de acuerdo a su naturaleza en cinco grupos [21, 33]:

- *Restricciones unarias*, aquellas que involucran un sólo evento, como por ejemplo, las clases de un curso no pueden ser programadas un día lunes.
- *Restricciones binarias*, aquellas que involucran dos eventos. Un ejemplo típico son las restricciones de topes de horarios para un curso que requiere un mismo recurso: profesor, sala de clases, etc.
- *Restricciones de capacidad*, como por ejemplo, las que se imponen al asignar cursos a salas de clase con capacidad suficiente.
- *Restricciones de separación de eventos*, aquellas que requieren que las actividades estén separadas o siguiendo algún patrón en el tiempo. Algunos ejemplos son las impuestas por políticas de la institución de respetar asignaciones de horarios en patrones predefinidos o las condiciones de no existencia de horas intermedias vacías.

- *Restricciones asociadas a los agentes*, como son las limitaciones en los horarios asignados para cumplir con las preferencias de los profesores.

Cuando todas las condiciones no pueden ser satisfechas simultáneamente, lo común es dividirlas en *requerimientos fuertes* que deben cumplirse obligatoriamente y *requerimientos suaves* que no son obligatorios pero sí deseables. La *calidad* de la programación obtenida dependerá del grado de cumplimiento de estas condiciones. Cuando el problema es enfrentado mediante el uso de modelos de optimización, los requerimientos fuertes son utilizados como restricciones, mientras que los requerimientos blandos son incluidos como un término en la función objetivo, las cuales al ser violadas se penalizan.

Entre las diversas técnicas de modelamiento y solución consideradas en la literatura, se destacan el uso de modelos de optimización, principalmente programación lineal y programación lineal entera, heurísticas y metaheurísticas, además de la programación por restricciones o *constraint programming*. Entre los trabajos que consideran este último enfoque podemos citar [1, 30, 43].

Numerosos son los trabajos que reportan el uso de heurísticas y metaheurísticas [26, 33, 39, 42]. Otros trabajos se concentran en el uso de metodologías particulares como por ejemplo: búsquedas tabú [4], *simulated annealing* [26], colonias de hormigas [46], algoritmos genéticos [50], búsquedas locales [29], métodos multiagentes [37, 48], metaheurísticas [14, 15, 41]. Algunos trabajos, como [20, 51], analizan métodos híbridos que combinan una o más metodologías básicas enunciadas anteriormente.

La metodología presentada en este trabajo se basa en un modelo de programación entera. Hay una gran cantidad de trabajos que atacan problemas similares utilizando modelos de optimización.

Al-Yacoob y Sherali [2] proponen dos modelos para la asignación de profesores a las clases. Estos toman como información de entrada la asignación previa de clases a horarios y tratan de asignar los profesores de manera a satisfacer sus preferencias. Uno de los modelos realiza la asignación sin modificar los horarios preestablecidos, mientras que el otro considera la posibilidad de realizar modificaciones en los horarios manteniendo el uso eficiente de las salas de clase. En otro trabajo [3], estos mismos autores, proponen el uso de un modelo de programación entera mixta que tiene como objetivo mejorar las programaciones de clases e incorporar nuevas condiciones impuestas por la *Kuwait University*.

Avella y Vasil'Ev [6] proponen un algoritmo de tipo *branch-and-cut* para un problema de asignación de horarios. La formulación utilizada está basada en un problema de *set packing* y son utilizadas como cortes desigualdades del tipo cliques. El problema es estudiado poliedralmente y son derivadas otras familias de desigualdades, no válidas en general para el *set packing*, y sus problemas de separación.

Daskalaki *et al.* [23] proponen un modelo de programación entera para un problema de generación de horarios. El modelo considera como función objetivo una función lineal que representa las preferencias de los profesores en horarios y en salas asignadas para sus clases. Los autores también analizan cómo la definición apropiada de los coeficientes en esta función objetivo permite reducir el espacio de soluciones y vuelve tratable el problema. Daskalaki y Birbas [22] presentan un algoritmo en dos etapas para la solución de un modelo de programación entera para la programación de horarios. En una primera etapa, el modelo de programación entera es relajado, eliminándose restricciones que corresponden a la contigüidad de sesiones para algunos cursos que así lo requieren. Estas condiciones son recuperadas en la segunda etapa, donde los horarios diarios son optimizados, incluyendo las restricciones no consideradas en la primera etapa.

MirHassani [35] presenta un modelo de programación entera para un problema similar al considerado en nuestro trabajo. El modelo propuesto asigna cada clase individualmente a un horario. Junto con las restricciones de topes horarios y usos de salas, son impuestas condiciones y políticas de la institución, como por ejemplo, que las clases de los cursos que tienen más de una clase por semana no pueden ser programadas el mismo día o en días consecutivos o la existencia de combinaciones de horarios que permitan tomar a un alumno todos los cursos de un semestre de su carrera, de acuerdo al plan de estudios.

Todos estos trabajos presentan modelos que asignan por separado cada clase de un curso a un bloque horario y por medio de restricciones se imponen las condiciones de regularidad de las clases. A diferencia de estos trabajos, Qualizza y Serafini [40] proponen un modelo de programación entera basado en generación de columnas. En este caso, las columnas están asociadas al horario semanal de un curso, es decir, todas las clases de la semana son programadas simultáneamente. Las restricciones en el problema maestro corresponden a la ocupación de salas y a los topes horarios. Las restricciones particulares a cada curso están incluidas en el subproblema.

El modelo considerado en nuestro trabajo está basado en una idea similar. Ciertos patrones de horarios son predefinidos y los cursos son asignados a estos patrones. A diferencia de lo propuesto en [40], las variables son definidas explícitamente y no generadas dinámicamente. Otra diferencia importante es que en nuestra propuesta, las clases auxiliares y las clases de cátedra de un curso son programadas por separado.

Existen otros trabajos que basados en modelos de programación entera presentan particularidades, como el uso de múltiples criterios o enfoques en varias etapas. Badri [7] propone un modelo de programación entera multiobjetivo para la programación de clases. El autor analiza un procedimiento en dos etapas para resolver el modelo propuesto. En una primera etapa, se busca maximizar las preferencias de los profesores en relación a los cursos

que les son asignados y, en la segunda, se asignan los horarios de clases a los cursos buscándose maximizar las preferencias de los profesores en cuanto a sus horarios de clases. Badri *et al.* [8] proponen un modelo de programación entera multiobjetivo. Este nuevo modelo busca asignar, simultáneamente, profesores a cursos y diseñar el horario de estos cursos de manera de maximizar las preferencias de los profesores. Estas preferencias, tanto por horarios como por cursos, son incluídas por medio de una tabla que registra las tres opciones preferidas tanto de cursos como de horarios. Dimopoulos y Miliotis [25] analizan la implementación de un sistema computacional de generación de horarios en un ambiente distribuido utilizando un modelo de programación entera para diseñar los horarios para cada departamento. El sistema cuenta con una base de datos central y un procedimiento automático se encarga de generar los problemas de optimización para cada departamento y de resolver los conflictos que pudieran aparecer. Stallaert [47] propone una metodología que divide el problema en dos subproblemas: primero se programan los cursos principales mediante un modelo de programación entera y, luego, utilizando esta programación como información de entrada, se programa el resto de los cursos resolviendo una variante de un problema de asignación cuadrático. Tripathy [49] propone un modelo de programación entera para un problema de programación de clases. El autor incorpora una simplificación al problema, agrupando cursos que pueden ser programados en el mismo horario. El trabajo considera un algoritmo basado en relajación lagrangiana y un algoritmo de tipo *branch-and-bound* para resolver la problemática propuesta.

Finalmente, algunos trabajos se concentran en el desarrollo y construcción de sistemas de apoyo a las decisiones, como por ejemplo: [4, 18, 24, 28, 31, 32, 36, 37, 38]. En particular, McCollum [34] analiza cuestiones prácticas y problemas que surgieron a la hora de implementar un sistema centralizado en una universidad británica.

2.2. Descripción de Problema

La Facultad de Ingeniería imparte en total cuatro carreras en pregrado: Ingeniería Civil Industrial, Ingeniería Civil en Computación, Ingeniería Civil en Obras Civiles e Ingeniería en Construcción. Las primeras tres carreras tienen una duración de 12 semestres, en tanto que la última, tiene una duración de 10 semestres. Cabe destacar que existe una malla curricular, la cual determina el orden en que los alumnos deben tomar los distintos cursos para cada carrera. De este modo, es posible caracterizar a cada curso por el semestre en que se ubica dentro del plan de estudio de la carrera. Esta información es útil a la hora de definir qué cursos no pueden ser dictados en forma simultánea en un mismo bloque horario.

Cada semestre se dictan en promedio 150 cursos los que tienen un número variable de secciones. Para un curso el número de secciones puede variar desde una sección a diez secciones paralelas. Los cursos tienen dos tipos de clases: *cátedra* y *auxiliares*. Cada curso tiene al menos una clase de *cátedra* semanalmente con un máximo de tres, en cambio para el caso de las clases auxiliares, existe la posibilidad de que un curso no tenga ninguna clase auxiliar o, en su defecto, tenga a lo más una clase por semana. Las clases se realizan de lunes a viernes en bloques de 1 hora y media de duración. Cada día se compone de 6 bloques horarios definidos por las letras A, B, C, D, E y F². Las clases de *cátedra* deben seguir un patrón horario³ definido por la Facultad. Respecto a las clases *auxiliares*, la Facultad tiene como condición deseable que estas se realicen el día miércoles, de no ser posible pueden dictarse en cualquier día y bloque horario. Cabe destacar que una de las condiciones impuestas es que para cada curso se debe respetar el mismo patrón horario para cada semana del semestre.

La Facultad de Ingeniería cuenta con 45 salas de clase, las cuales son compartidas por todos los cursos de las carreras impartidas. Las salas de clase se caracterizan por su capacidad, definida como el número máximo de alumnos que es posible asignar para un bloque horario. Estas salas se clasifican en 6 grupos: salas normales, laboratorios de física, laboratorios de computación, laboratorios de obras civiles, laboratorios de simulación de procesos y un Auditorio. El Auditorio es una sala de clase de mayor capacidad y tecnología utilizado para eventos importantes y de alta convocatoria. Cabe destacar que las salas de clase son el recurso escaso de la Facultad.

La Facultad dispone de un staff de 150 profesores para realizar las clases de *cátedra* y cerca de 100 profesores *auxiliares* para realizar las clases *auxiliares*. Cada profesor de *cátedra* se caracteriza por los cursos que dicta y por su disposición horaria. Para el caso de los profesores *auxiliares* no se consideran sus disposiciones, ya que son alumnos de la Facultad quienes, en estricto rigor, se acomodan a los horarios preestablecidos por la Facultad para los distintos cursos.

Actualmente, la programación horaria es generada por un equipo conformado por 3 profesionales, los cuales demoran en promedio un mes para obtener la programación final. La programación final para un semestre cualquiera se genera principalmente sobre la base de la programación horaria utilizada en el semestre anterior. Esta última es actualizada solamente al existir un nuevo requerimiento, al incorporar un curso nuevo en el plan de estudios o al haber cambios en las preferencias horarias de los profesores. Cabe destacar que esta programación no está exenta de errores, observándose una serie de ineficien-

²El bloque A es el primer bloque del día y el bloque F es el último.

³Se entiende como patrón horario a una combinación entre uno o más días con uno o más bloques horarios.

cias a la hora de realizar la asignación de salas de clase y un sinnúmero de conflictos entre los horarios de cursos de un mismo semestre.

Considerando los antecedentes planteados anteriormente, la generación de la programación de horarios y asignación de salas de clase se transforma en una tarea en extremo compleja y que consume una enorme cantidad de recursos. Por este motivo, el modelo propuesto busca que los requerimientos impuestos por la Facultad sean apropiadamente estructurados mediante la formulación de un modelo de programación entera. La resolución de este modelo entregará la programación de horarios y asignación de salas de clase de manera óptima respecto de alguna función objetivo.

2.2.1. Requerimientos Impuestos por la Facultad de Ingeniería

A continuación, se describen los requerimientos impuestos por la Facultad de Ingeniería para la programación horaria y asignación de salas de clase. Estos requerimientos fueron categorizados en dos grupos: *fuertes* y *suaves*. Los requerimientos *fuertes* deben ser cumplidos obligatoriamente y los requerimientos *suaves*, que si bien no son obligatorios, representan condiciones deseables para la Facultad. Los requerimientos *suaves* se incorporaron dentro de la función objetivo, la cual trata de minimizar el número de veces en que no se cumplen estos requerimientos. Cada vez que estos no se cumplan se incurre en un penalización.

Requerimientos Fuertes

1. Cada curso debe ser asignados a una sala de clase con capacidad suficiente para la demanda estimada de alumnos para dicho curso.
2. En una sala de clase, en un mismo día y bloque horario, se puede realizar a lo más una clase (*cátedra* o *auxiliar*).
3. Un profesor no puede dictar más de una clase a la vez.
4. Se deben respetar los horarios disponibles de los profesores.
5. No deben existir topes de horarios entre cursos de un mismo semestre.
6. Cada curso debe seguir alguno de los patrones horarios impuestos por la Facultad para la realización de sus clases. Existen cuatro tipos de patrones horarios:
 - a) *Patrón 1*: Está compuesto por la combinación entre un día de la semana y un bloque horario. Por ejemplo, un patrón de esta clase es {LU-A} que corresponde al día lunes bloque horario A.

- b) *Patrón 2:* Está compuesto por la combinación de un día de la semana y dos bloques horarios consecutivos. Por ejemplo un patrón de este grupo puede ser: {MA-C, MA-D} que corresponde al patrón que contiene el día martes y los bloques horarios C y D.
- c) *Patrón 3:* En este caso los patrones deben contener dos días de la semana distintos y un bloque horario. Las combinaciones posibles son: lunes-jueves y martes-viernes. Un ejemplo de este tipo de patrón es {LU-E, JU-E}, o sea el patrón contiene los días lunes y jueves en el bloque horario E.
- d) *Patrón 4:* Este grupo de patrones está compuesto por la combinación de tres días de la semana distintos y un bloque horario. En este caso, se permite cualquier combinación de días, es decir, los patrones de este tipo son aquellos de la forma {D1-X, D2-X, D3-X} donde $D1$, $D2$ y $D3$ son tres días distintos de la semana y X es un bloque horario cualquiera.

Requerimientos Suaves

1. Las clases auxiliares deben realizarse de preferencia los días miércoles en cualquier bloque horario. De no ser posible esta asignación, se pueden realizar en cualquier día y bloque horario.
2. Se debe evitar, en lo posible, asignar cursos al Auditorio.

En la próxima sección se detalla el modelo propuesto para enfrentar el problema descrito.

3. Enfoque de Solución

Se considera como unidad básica de modelamiento a lo que llamamos un (par) *curso-sección*. Esta unidad corresponde simplemente a una sección de un curso. La metodología desarrollada se basa en un modelo de programación entera que integra la definición de la programación de horarios y la asignación de salas de clase. El modelo, de manera similar al propuesto por Qualizza y Serafini [40], asigna de manera conjunta todas las clases de cátedra o auxiliares de un curso-sección correspondientes a una semana.

Para esto, se definen combinaciones de bloques horarios-salas factibles. Estas combinaciones de horarios y salas corresponden a agrupaciones de pares (bloque horario, sala) que llamamos *patrones horarios-salas* o simplemente, *patrones HS*. La idea es que el modelo decida si programar o no las clases de cátedra o auxiliares de un curso-sección directamente en un patrón HS sin

identificar cada una de las clases de un curso-sección en una variable distinta, sino que asignándolas todas como conjunto al patrón. Por ejemplo, si un curso posee dos cátedras que deben ser asignadas a bloques horarios consecutivos del mismo día y a la misma sala, entonces sus dos cátedras se asignarán a algún patrón HS de la forma: $\{(t, s), (t + 1, s)\}$ con t y $t + 1$ bloques horarios consecutivos del mismo día y s una sala de clase adecuada para el curso-sección en cuestión.

De esta manera, las condiciones que todas las clases de cátedra de un curso-sección deberán ser realizadas en la misma sala y que los bloques horarios asignados correspondan a una de las combinaciones definidas por la Facultad se manejan de manera implícita al definir las variables.

Antes de describir el modelo, definimos la notación utilizada. Sea I el conjunto de pares curso-sección, S el conjunto de salas disponibles, P , el conjunto de profesores y H el conjunto de semestres de las mallas académicas de las carreras impartidas por la Facultad. El conjunto de bloques horarios de la semana es denotado por T . Los elementos de este conjunto son pares de la forma “día-bloque”, por ejemplo LU-A es el primer bloque de la semana y MI-F, es el último bloque del día miércoles. El conjunto de los patrones HS es denotado por B . Por último, existen ciertos grupos de cursos cuyas clases no deben topar por diferentes motivos (por ejemplo, porque corresponden al mismo semestre del plan de estudios). A la familia de grupos de curso-sección cuyas clases no deben topar en horario lo denotamos por L .

Adicionalmente, son necesarios los siguientes subconjuntos de los conjuntos que acabamos de definir:

- $CPR_p \subseteq I$: Conjunto de cursos-sección asignados profesor p .
- $SMC_{hl} \subseteq I$: l -ésimo grupo de cursos-sección del semestre h que no deben tener topes horarios.
- $HR_p \subseteq T$: Conjunto de disponibilidad horaria del profesor p .
- $MIE \subseteq B$: Conjunto de patrones HS que poseen bloques horarios del día miércoles.
- $PC_i \subseteq B$: Conjunto de patrones HS que pueden ser utilizados por las cátedras del curso-sección i .
- $PA_i \subseteq B$: Conjunto de patrones HS que pueden ser utilizados por las clases auxiliares del curso-sección i .
- $PSC_{si} \subseteq B$: Conjunto de patrones HS que ocupan la sala s y pueden ser utilizados por las cátedras del curso-sección i .
- $PSA_{si} \subseteq B$: Conjunto de patrones HS que ocupan la sala s y pueden ser utilizados por las clases auxiliares del curso-sección i .

- $PTC_{ti} \subseteq B$: Conjunto de patrones HS que contienen al bloque horario t y pueden ser utilizados por las cátedras del curso-sección i .
- $PTA_{ti} \subseteq B$: Conjunto de patrones HS que contienen al bloque horario t y pueden ser utilizados por las clases auxiliares del curso-sección i .
- $PSTC_{sti} \subseteq B$: Conjunto de patrones HS que ocupan la sala s y el bloque horario t y pueden ser utilizados por las cátedras del curso-sección i .
- $PSTA_{sti} \subseteq B$: Conjunto de patrones HS que ocupan la sala s y el bloque horario t y pueden ser utilizados por las clases auxiliares del curso-sección i .
- $AX \subseteq I$: Conjunto de cursos-sección que poseen clases auxiliares.

Se definen dos conjuntos de variables binarias asociadas a la asignación de clases de cátedra y clases auxiliares a patrones HS, respectivamente. Estas variables son:

$$x_{ib} = \begin{cases} 1 & \text{si las clases de cátedra del curso-sección } i \text{ se realizan de acuerdo} \\ & \text{al patrón } b \in PC_i, \\ 0 & \text{en caso contrario;} \end{cases}$$

y

$$y_{ib} = \begin{cases} 1 & \text{si las clases auxiliares del curso-sección } i \text{ se realizan de acuerdo} \\ & \text{al patrón } b \in PA_i, \\ 0 & \text{en caso contrario;} \end{cases}$$

Con la notación definida, el modelo de programación entera es el siguiente:

$$\text{minimizar } z = \sum_{i \in I} \left(\sum_{b \in L \setminus MIE} y_{ib} + \sum_{b \in PSA_{AU_i}} y_{ib} \right) + \sum_{i \in I} \sum_{b \in PSC_{AU_i}} x_{ib}$$

sujeto a

$$\sum_{b \in PC_i} x_{ib} = 1 \quad \text{para todo } i \in I, \quad (1)$$

$$\sum_{b \in PA_i} y_{ib} = 1 \quad \text{para todo } i \in AX, \quad (2)$$

$$\sum_{i \in I} \sum_{b \in PSTC_{sti}} x_{ib} + \sum_{i \in I} \sum_{b \in PSTA_{sti}} y_{ib} \leq 1 \quad \text{para todo } s \in S, t \in T, \quad (3)$$

$$\sum_{b \in PTC_{ti}} x_{ib} + \sum_{b \in PTA_{ti}} y_{ib} \leq 1 \quad \text{para todo } i \in I, t \in T, \quad (4)$$

$$\sum_{i \in CPR_p} \sum_{b \in PTC_{ti}} x_{ib} \leq 1 \quad \text{para todo } p \in P, t \in HR_p, \quad (5)$$

$$\sum_{i \in SMC_{hl}} \sum_{b \in PTC_{ti}} x_{ib} + \sum_{i \in SMC_{hl}} \sum_{b \in PTA_{ti}} y_{ib} \leq 1 \quad \text{para todo } h \in H, l \in L, t \in T, \quad (6)$$

$$x_{ib} \in \{0, 1\} \quad \text{para todo } i \in I, b \in PC_i, \quad (7)$$

$$y_{ib} \in \{0, 1\} \quad \text{para todo } i \in AX, b \in PA_i. \quad (8)$$

La función objetivo representa la minimización de clases auxiliares asignadas a bloques horarios que no pertenecen al día miércoles más la cantidad de clases asignadas al Auditorio de la Facultad (AU).

Las restricciones (1) y (2) garantizan que para todos los cursos se asignen patrones que programen todas las clases de cátedra y auxiliares. La restricción (3) impide la asignación de más de un curso-sección a la misma sala en el mismo horario. La restricción (4) imposibilita que las clases de cátedra y auxiliares de un mismo curso sean programadas en el mismo horario. La restricción (5) evita que se programen en el mismo horario clases dictadas por el mismo profesor. Esta restricción no garantiza que las clases sean asignadas a bloques en que el profesor está disponible. Esto es controlado al definir el subconjunto PC_i de patrones HS que pueden ser utilizados para dictar las clases de cátedra del curso-sección i . Finalmente, la restricción (6) controla que no se produzcan topes de horarios entre clases de cursos diferentes que así lo requieran. En el caso particular de la aplicación considerada, se evitaron los topes entre cursos del mismo semestre en el plan de estudios de alguna de las carreras.

4. Resultados Experimentales

En esta sección se realiza una comparación de los resultados obtenidos entre el modelo propuesto y el sistema manual que actualmente utiliza la Facultad, para la generación de horarios y asignación de salas de clase. El objetivo central de esta sección es mostrar las ventajas de la utilización del modelo

matemático. Para esto se comparan distintos indicadores de desempeño. Estos indicadores miden el cumplimiento de las condiciones deseables y los objetivos de la Facultad.

Los indicadores utilizados fueron los siguientes:

I_1 : Porcentaje de clases auxiliares el día miércoles.

I_2 : Número de cursos asignados al Auditorio.

I_3 : Número cursos con horarios de cátedra sin patrón

I_4 : Número cursos asignados a salas de clase distintas para la realización de sus actividades.

I_5 : Número cursos asignados a salas con capacidad insuficiente.

I_6 : Número de pares de cursos asignados a una misma sala de clases.

I_7 : Número de semestres sin horarios factibles.

I_8 : Tiempo de generación de la programación.

El indicador I_3 se refiere a la existencia de cursos que no son dictados en los patrones definidos por la Facultad. El indicador I_7 cuenta los semestres que no tienen un horario factible. Esto quiere decir que, al considerar todos los cursos que componen a dicho semestre de acuerdo al plan de estudios, no existe una combinación de horarios que permita inscribirlos sin topes de horario.

4.1. Descripción de la Instancia

La instancia utilizada para la comparación de resultados fue suministrada por la Facultad de Ingeniería. Esta instancia está compuesta por 200 cursos, 150 profesores y 45 salas de clase.

El modelo de programación entera tiene 284.766 variables de decisión y 13.185 restricciones. Este modelo fue modelado utilizando la herramienta GAMS 22.5 y resuelto mediante el solver CPLEX 10.0 en un ordenador con procesador Intel Centrino Duo de 1.83 GHZ con 2 GB de memoria RAM.

4.2. Comparación de Resultados

La solución entregada por el sistema actual no cumple con las condiciones deseables para la operatividad normal de la institución. La Tabla 4.2 muestra los resultados obtenidos para la instancia descrita anteriormente.

Los dos primeros indicadores (I_1 y I_2) están relacionados directamente con la función objetivo del modelo. Existe una mejora significativa en la asignación de clases auxiliares para los días miércoles, pasando de un 43 % en el sistema

Indicadores	Sistema Actual	Modelo Propuesto
I_1	43 %	92 %
I_2	0	1
I_3	2	0
I_4	60	0
I_5	105	0
I_6	10	0
I_7	2	0
I_8	1 mes	2 min.

Cuadro 1: Comparación de resultados para las programaciones generadas por el modelo propuesto y el sistema actual.

actual a un 92 % obtenido por el modelo propuesto considerando el total de clases. Cabe destacar que este requerimiento nunca se podrá alcanzar en un 100 % debido a que la cantidad de salas de clase de la Facultad no puede albergar a todos los cursos que tienen clases auxiliares en un día. Respecto del segundo indicador, a pesar de que el sistema manual no genera asignaciones de cursos en el Auditorio, podemos señalar que trajo consigo que algunos cursos fueran asignados a salas de clase con capacidad insuficiente, ya que este Auditorio es la sala de clase de mayor capacidad.

Respecto a la condiciones deseables para la operatividad del sistema, el modelo propuesto asigna todos los cursos a un patrón establecido, no permite que un curso sea asignado a salas de clase diferentes, no permite la asignación de dos o más cursos a una misma sala de clase en un bloque horario y genera siempre al menos una combinación factible para los cursos de un semestre particular.

El sistema actual viola estas condiciones, ya que se obtienen 60 cursos asignados en salas de clase diferentes y 105 cursos asignados a una sala con capacidad insuficiente y asigna más de un curso a una misma sala de clases. Este tipo de ineficiencias causa problemas dentro del alumnado y, en especial, cuando no se respeta la capacidad de las salas de clase. Esto último en la práctica ocasiona que un número significativo de alumnos no pueda asistir a esos cursos.

Otro punto a destacar es que en la instancia estudiada para dos semestres de los planes de estudio no existe un horario factible para los distintos cursos que los componen. Esto significa que un alumno que curse dichos semestres, no podrá seleccionar todos sus cursos o, en su defecto, siempre tendrá algún tope de horario. El modelo propuesto proporciona siempre un horario factible permitiendo al alumnado inscribir todos los cursos correspondientes a un semestre específico.

Finalmente un indicador muy importante a la hora de definir la programación de cursos y la asignación de salas de clase es el tiempo que demora dicho proceso. Las mejoras en este sentido son sustanciales, pasando desde un mes para el procedimiento manual a dos minutos con el modelo propuesto. Esta mejora permite no solo obtener una solución óptima en un corto tiempo, sino que es posible explorar múltiples escenarios, obtener soluciones alternativas y poder reaccionar ante eventos inesperados, como lo son la incorporación de cursos nuevos sobre la marcha.

En resumen, los beneficios reportados por el modelo propuesto son evidentes. La utilización de este modelo permitirá a la Facultad de Ingeniería eliminar todos los conflictos y, satisfacer todas las condiciones de operación, y además, ser eficiente en la asignación de un recurso escaso como lo son las salas de clase.

5. Conclusiones

Este trabajo presenta un modelo de programación entera para la programación de horarios y asignación de salas de clase para la Facultad de Ingeniería de la Universidad Diego Portales. Este modelo asigna simultáneamente todas las clases de cátedra o auxiliares de un curso a algún patrón horario-sala. Este enfoque es semejante al propuesto por Qualizza y Serafini [40] y se diferencia de otros modelos presentados en la literatura que asignan cada clase a bloques horarios por separado y fuerzan la formación de patrones a través de restricciones.

Considerando el tamaño del problema, la tarea de generar la programación de cursos manualmente se transforma en una labor en extremo compleja y sujeta a múltiples errores. La utilización del modelo propuesto garantiza la satisfacción de todos los requerimientos obligatorios y mejora en forma significativa el cumplimiento de las condiciones deseables. En particular, mientras la programación actual sólo consigue programar alrededor de un 40 % de las clases auxiliares el día miércoles, la solución obtenida por el modelo permite hacerlo con más del 90 %. Adicionalmente, se reduce sustancialmente el tiempo requerido para la obtención de la programación de horarios y asignación de salas de clase.

Actualmente, se está desarrollando un sistema computacional que utiliza como optimizador el modelo de programación entera propuesto e incorporará diferentes opciones mediante una interfaz amigable. Con la implementación del sistema computacional será posible automatizar el proceso de generación de la programación de horarios y asignación de salas aumentando la calidad de las programaciones, la flexibilidad ante cambios en los requerimientos y condiciones deseables y la posibilidad de explorar múltiples escenarios.

Una línea interesante para profundizar en el trabajo de investigación es la adaptación del modelo propuesto para su resolución mediante generación de columnas y la implementación de un algoritmo tipo *branch-and-price*.

Agradecimientos: Al Instituto Científico Milenio “Sistemas Complejos de Ingeniería” P04-066-F por el apoyo económico para la concreción de este proyecto. Adicionalmente, los autores agradecen la buena disposición del decano de la Facultad de Ingeniería de la Universidad Diego Portales, José Manuel Robles, al secretario de estudios de la Escuela de Ingeniería Industrial, Claudio Gutiérrez y, de manera especial, a Francia Lara I. por sus valiosos comentarios para la edición final de este documento.

Referencias

- [1] S. Abdennadher and M. Marte. University course timetabling using constraint handling rules. *Applied Artificial Intelligence*, 14:311–325, 2000.
- [2] S. Al-Yakoob and S. Sheralli. Mathematical programming models and algorithms for a class-faculty assignment problem. *European Journal of Operational Research*, 173:488–507, 2006.
- [3] S. Al-Yakoob and S. Sheralli. A mixed-integer programming approach to a class timetabling problem: A case study with gender policies and traffic considerations. *European Journal of Operational Research*, 180:1028–1044, 2007.
- [4] R. Alvarez-Valdes, E. Crespo, and J. Tamarit. Design and implementation of a course scheduling system using tabu search. *European Journal of Operational Research*, 137:512–523, 2002.
- [5] A. Asratian and D. de Werra. A generalized class-teacher model for some timetabling problems. *European Journal of Operational Research*, 143:531–542, 2002.
- [6] P. Avella and L. Vasil’ev. A computational study of a cutting plane algorithm for university course timetabling. *Journal of Scheduling*, 8(6):497–514, 2004.
- [7] M. Badri. A two-stage multiobjective scheduling model for [faculty-course-time] assignments. *European Journal of Operational Research*, 94:16–28, 1996.

- [8] M. Badri, D.L. Davis, D.F. Davis, and J. Hollingsworth. A multi-objective course scheduling model: Combining faculty preferences for courses and times. *Computers and Operations Research*, 25:303–316, 1998.
- [9] N. Boland, B. Hughes, L. Merlot, and P. Stuckey. New integer linear programming approaches for course timetabling. *Computers and Operations Research*, 35:2209–2233, 2008.
- [10] E. Burke and M. Carter, editors. *The Practice and Theory of Automated Timetabling II: Selected Papers from the 2nd International Conference on the Practice and Theory of Automated Timetabling*. LNCS 1408, Springer-Verlag, 1998.
- [11] E. Burke and P. De Causmaecker, editors. *The Practice and Theory of Automated Timetabling IV: Selected papers from the 4th International Conference (PATAT IV)*. LNCS 2740, Springer-Verlag, 2003.
- [12] E. Burke and W. Erben, editors. *Practice and Theory of Automated Timetabling: Selected papers from the 3rd International Conference*. LNCS 2079, Springer-Verlag, 2001.
- [13] E. Burke, K. Jackson, J.H. Kingston, and R. Weare. Automated university timetabling: The state of the art. *Computer Journal*, 40(9):565–571, 1997.
- [14] E. Burke, B. MacCarthy, S. Petrovic, and R. Qu. Knowledge discovery in a hyper-heuristic for course timetabling using case-based reasoning. In Burke and Causmaecker [11], pages 276–287.
- [15] E. Burke, B. McCollum, A. Meisels, S. Petrovic, and R. Qu. A graph-based hyper-heuristic for educational timetabling problems. *European Journal of Operational Research*, 176:177–192, 2007.
- [16] E. Burke and H. Rudová, editors. *PATAT 2006: Proceedings of The 6th International Conference on the Practice and Theory of Automated Timetabling*, Masaryk University, Brno, República Checa, 2006. Disponible en la página http://patat06.muni.cz/doc/PATAT_2006_Proceedings.pdf.
- [17] E. Burke and M. Trick, editors. *The Practice and Theory of Automated Timetabling V: Selected Revised Papers from the 5th International Conference on the Practice and Theory of Automated Timetabling*. LNCS 3616, Springer-Verlag, 2005.
- [18] M. Carter. A comprehensive course timetabling and student scheduling system at the university of waterloo. In Burke and Erben [12], pages 64–82.

- [19] L. C. Chambers, editor. *The Practical Handbook of Genetic Algorithms*, volume 1. CRC Press, 1995.
- [20] M. Chiarandini, M. Birattari, K. Socha, and O. Rossi-Doria. An effective hybrid algorithm for university course scheduling. *Journal of Scheduling*, 9:403–432, 2006.
- [21] D. Corne, P. Ross, and H. Fang. *Evolving timetables*, pages 219–276. Volume 1 of Chambers [19], 1995.
- [22] S. Daskalaki and T. Birbas. Efficient solutions for a university timetabling problem through integer programming. *European Journal of Operational Research*, 160:106–120, 2005.
- [23] S. Daskalaki, T. Birbas, and E. Housos. An integer programming formulation for a case study in university timetabling. *European Journal of Operational Research*, 153:117–135, 2004.
- [24] M. Dimopoulou and P. Miliotis. Implementation of a university course and examination timetabling system. *European Journal of Operational Research*, 130:202–213, 2001.
- [25] M. Dimopoulou and P. Miliotis. An automated university course timetabling system developed in a distributed environment: A case study. *European Journal of Operational Research*, 153:136–147, 2004.
- [26] M. Elmohamed, P. Coddington, and G. Fox. A comparison of annealing techniques for academic course scheduling. In Burke and Carter [10], pages 92–112.
- [27] J. Ferland and C. Fleurent. Saphir - a decision-support system for course scheduling. *Interfaces*, 24:105–115, 1994.
- [28] L. Foulds and D. Jonhson. Slotmanager: a microcomputer-based decision support system for university timetabling. *Decision Support Systems*, 27:367–381, 2000.
- [29] L. Di Gaspero and A. Schaerf. Multi-neighbourhood local search with application to course timetabling. In Burke and Causmaecker [11], pages 262–275.
- [30] H. Goltz and D. Matzke. University timetabling usin constraint logic programming. In *Practical Aspects of Declarative Languages*, volume 1551 of *LNCS*, pages 320–334. Springer-Verlag, 1999.
- [31] K. Haase, H. Scheel, and D. Sebastian. Management of lecture-rooms. model, method and intenet-based application for efficient course scheduling. *Wirtschaftsinformatik*, 46:87–95, 2004.

- [32] T. Hinkin and G. Thompson. Scheduleexpert: Scheduling courses at cornell university school of hotel administration. *Interfaces*, 32:45–57, 2002.
- [33] R. Lewis. A survey of metaheuristics-based techniques for university timetabling problems. *OR Spectrum*, 30:167–190, 2008.
- [34] B. McCollum. The implementation of a central timetabling system in a large british civic university. In Burke and Carter [10], pages 237–253.
- [35] S. MirHassani. A computational approach to enhancing course timetabling with integer programming. *Applied Mathematics and Computation*, 175:814–822, 2006.
- [36] E. Mooney, R. Rardin, and W. Parmenter. Large-scale classroom scheduling. *IIE Transactions*, 28:369–378, 1996.
- [37] M. Oprea. Mas_up_uct: A multi-agent system for university course timetabling scheduling. *International Journal of Computer Communications and Control*, 2:94–102, 2007.
- [38] B. Paetcher, R. Rankin, and A. Cumming. Improving a lecture timetabling system for university wide-use. In Burke and Carter [10], pages 156–165.
- [39] P. Pongcharoen, W. Promtet, P. Yenradee, and C. Hicks. Stochastic optimisation timetabling tool for university course scheduling. *International Journal of Production Economics*, 112:903–918, 2008.
- [40] A. Qualizza and P. Serafini. A column generation scheme for faculty timetabling. In Burke and Trick [17], pages 161–173.
- [41] P. Rattadilok, A. Gaw, and R. Kwan. Distributed choice function hyper-heuristic for timetabling and scheduling. In Burke and Trick [17], pages 51–67.
- [42] O. Rossi-Doria, M. Sampels, M. Birattari, M. Chiarandini, M. Dorigo, L. Gambardella, J. Knowles, M. Manfrin, M. Mastrolilli, B. Paetcher, L. Paquete, and T. Stutzle. A comparison of the performance of different metaheuristics on the timetabling problem. In Burke and Causmaecker [11], pages 329–351.
- [43] H. Rudova and K. Murray. University course timetabling with soft constraints. In Burke and Causmaecker [11], pages 310–328.
- [44] K. Sandhu. *Automating class schedule generation in the context of a university timetabling information system*. PhD thesis, School of Management, Griffith University, 2001.

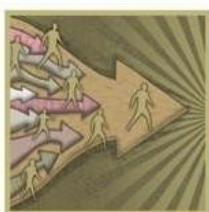
- [45] A. Schaefer. A survey of automated timetabling. *Artificial Intelligence Review*, 13:87–127, 1999.
- [46] K. Socha, M. Sampels, and M. Manfrin. Ant algorithm for the university course timetabling problem with regard to the state-of-the-art. In *Applications of evolutionary computing*, volume 2611 of *LNCS*, pages 334–345. Springer-Verlag, 2003.
- [47] J. Stallaert. Automated timetabling improves course scheduling at UCLA. *Interfaces*, 27(4):67–81, 1997.
- [48] D. Strnad and N. Guid. A multi-agent system for university timetabling. *Applied Artificial Intelligence*, 21:137–153, 2007.
- [49] A. Tripathy. School timetabling – A case in large binary integer linear programming. *Management Science*, 30(12):1473–1489, 1984.
- [50] Y. Wang. Usin genetic algorithm methods to solve course scheduling problems. *Expert Systems with Applications*, 25:39–50, 2003.
- [51] G. White and J. Zhang. Generating complete university timetables by combining tabu search with constraint logic. In Burke and Carter [10], pages 187–198.

Programas de Postgrado Impartidos por el DII



MGO

MAGISTER EN GESTIÓN DE OPERACIONES



MAYOR INFORMACIÓN

Domeyko 2313, Piso 1, Santiago de Chile
Teléfonos: (562) 9784017- (562) 9784073
Email: julie@diu.uchile.cl
www.dii.uchile.cl/mgo

El Magíster en Gestión de Operaciones busca formar profesionales de excelencia en esquemas de gestión, uso de modelos y tecnologías de información, con el propósito de que éstos desarrollen la capacidad de resolución de problemas complejos en el ámbito de la gestión de operaciones. El programa está reacreditado por CONAP.

El Magíster es impartido por el Centro de Gestión de Operaciones del Departamento de Ingeniería Industrial de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.

Calendario Académico

Semestre Otoño: Postulaciones: Octubre a 15 de Diciembre. Inicio de clases Marzo de cada año

Semestre Primavera: Postulaciones: Abril a 15 de Junio. Inicio de clases Julio de cada año.

Gestión de Operaciones

DOCTORADO en Sistemas de Ingeniería

www.sistemasdeingenieria.cl | doctorado@sistemasdeingenieria.cl | (56-2) 978 4017 - 978 4073

- Modelación de **sistemas y procesos** de alta complejidad.
- Técnicas avanzadas en **optimización y manejo** de datos.
- Cuerpo académico del **más alto nivel**.

UNIVERSIDAD
DE CHILE



Doctorado en Economía

La Facultad de Economía y Negocios y la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, a través de sus respectivos Departamentos de Economía e Ingeniería Industrial, ofrecen desde el año 2000 el primer programa de Doctorado en Economía de Chile. Este cuenta con el prestigio y la excelencia que otorga el sello de la Universidad que es líder en investigación y docencia a nivel país.

El Doctorado en Economía busca ser un referente regional en la formación de economistas de alto nivel y de esta forma atraer a estudiantes de excelencia graduados en Chile y América Latina, que tengan claros intereses académicos.

Nuestros egresados están capacitados para incorporarse a las diversas instituciones de investigación y de educación superior y a los grupos de investigación en instituciones públicas y privadas.

El cuerpo académico del Doctorado en economía es de primer nivel con profesores doctorados en las universidades de mayor prestigio y con una alta productividad académica, en términos de publicaciones.

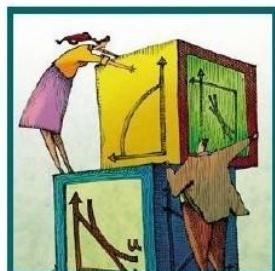


Plazos para postular (para inicio de clases en marzo del 2009)

Primer Proceso: Finaliza el segundo viernes de octubre. Se recomienda a quienes quieran postular a financiamiento interno y externo al programa enviar su postulación en esta fecha.

Segundo Proceso: Finaliza el primer viernes de diciembre. En este proceso se completarán los cupos que hayan quedado disponibles en el proceso anterior.

MAYOR INFORMACIÓN:
fmelis@diu.uchile.cl
562-9784073



El Magíster en Economía Aplicada (MAGEA) del Departamento de Ingeniería Industrial de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, busca formar profesionales de gran competencia analítica y una sólida base en economía.

El MAGEA es impartido por el Centro de Economía Aplicada (CEA) del Departamento de Ingeniería Industrial, líder en investigación en economía en Chile y en el desarrollo de propuestas para las políticas públicas. Adicionalmente, destacados académicos y profesionales son invitados a dictar clases en el programa. El programa está reacreditado por CONAP.

MAYOR INFORMACIÓN:
Domeyko 2313, Piso 1, Santiago de Chile
Teléfonos: (562) 9784084 - (562) 9784073
Email: magea@diu.uchile.cl
Página web: www.magea-uchile.cl

Requisitos de Admisión

Título Profesional, nacional o extranjero, que exija al menos 5 años de estudio o el grado de licenciado en campos disciplinarios afines a la especialidad.

Calendario Académico

Semestre Otoño: Postulaciones Octubre a 15 de Diciembre. Inicio de clases marzo de cada año.

Semestre Primavera: Postulaciones Abril a 15 de Junio. Inicio de clases julio de cada año.

MAGEA
MAGÍSTER EN ECONOMÍA APLICADA

