

REVISTA

Ingeniería de Sistemas

Volumen XXXV, Diciembre 2021

Predicción de arribos y programación de turnos en una estación de servicio

Octavio Eneas Biasoli / Javier Marengo

Inferencia Bayesiana de un Modelo Markoviano de Fútbol con aplicación en Scouting

Pablo Galaz / Sebastián Mena,
Denis Saure

Optimización de la zonificación y el ruteo de camiones en la recolección de residuos de un municipio del Gran Buenos Aires

Federico Bertero / Manuela Cerdeiro
Guillermo A. Durán / Nazareno A. Faillace

Remuneración de Redes de Transmisión en Chile por VNR: Principios, Problemas y Recomendaciones

Ronald Fischer / Rodrigo Moreno

R E V I S T A

INGENIERIA DE SISTEMAS

Volumen XXXV

Diciembre 2021

- Predicción de arribos y programación de turnos en una estación de servicio 5
Octavio Eneas Biasoli, Javier Marengo.
- Inferencia Bayesiana de un Modelo Markoviano de Fútbol con aplicación en Scouting 21
Pablo Galaz, Sebastián Mena, Denis Sauré.
- Optimización de la zonificación y el ruteo de camiones en la recolección de residuos de un municipio del Gran Buenos Aires 49
Federico Bertero, Manuela Cerdeiro, Guillermo A. Durán, Nazareno A. Faillace.
- Remuneración de Redes de Transmisión en Chile por Valor Nuevo de Reemplazo: Principios, Problemas y Recomendaciones 81
Ronald Fischer, Rodrigo Moreno.

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 Matemática e Instituto de Cálculo
FCEN-UBA y CONICET, Argentina
Departamento de Ingeniería Industrial
Universidad de Chile*

EDITOR ASOCIADO

Richard Weber

*Departamento de Ingeniería Industrial
Universidad de Chile*

AYUDANTE DE EDICIÓN

Cinthya Vergara

*Departamento de Ingeniería Industrial
Universidad de Chile*

COMITÉ EDITORIAL

René Caldentey

University of Chicago, Estados Unidos

Héctor Cancela

Universidad de la República, Uruguay

Rafael Epstein

Universidad de Chile, Chile

Luis Llanos

Westgate Energía, Chile

Javier Marengo

*Universidad Nacional de
General Sarmiento, Argentina*

Juan de Dios Ortúzar

*Pontificia Universidad
Católica de Chile, Chile*

Víctor Parada

Universidad de Santiago, Chile

Oscar Porto

GAPSO, Brasil

Lorena Pradenas

Universidad de Concepción, Chile

Nicolás Stier

Facebook Core Data Science, Estados Unidos

La Revista Ingeniería de Sistemas (RIS) es financiada parcialmente por el Instituto Sistemas Complejos de Ingeniería (ISCI), 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 las del Departamento de Ingeniería Industrial de la Universidad de Chile.

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

Representante legal: Richard Weber
Correo electrónico: ris@dii.uchile.cl
Diagramación y Diseño:

Dirección: Av. Beauchef 851, Santiago, Chile.
Web URL: www.dii.uchile.cl/ris
Cinthya Vergara

Carta Editorial Volumen XXXV

Nos es muy grato presentar este nuevo número de la Revista de Ingeniería de Sistemas (RIS) dedicado a temas de frontera en Investigación de Operaciones y disciplinas relacionadas. Queremos agradecer al Instituto Sistemas Complejos de Ingeniería (ISCI) por su colaboración para hacer posible esta publicación.

Este número contiene artículos de académicos y estudiantes de nuestro Departamento de Ingeniería Industrial, de investigadores del ISCI y de académicos y estudiantes de Universidades argentinas.

Nuestro objetivo a través de esta publicación es contribuir a la generación y difusión de las tecnologías modernas de investigación de operaciones, analítica, gestión y tecnología. La revista pretende destacar la importancia de generar conocimiento en estas áreas, orientado tanto a problemáticas nacionales como a la realidad de países de características similares de la región.

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

Richard Weber
Editor Asociado

Llamado a Presentar Trabajos

La Revista Ingeniería de Sistemas (RIS) busca constituir un canal de divulgación de los avances en las áreas de Gestión de Operaciones, Tecnologías de Información e Investigación Operativa, que incluya los mundos académico y empresarial. Son particularmente apropiados artículos orientados a la práctica de estas disciplinas, que estimulen su uso o den cuenta de aplicaciones innovadoras de ellas, especialmente en América Latina.

También son bienvenidos artículos con análisis del estado del arte en un campo particular y de la forma en que los avances en dicho campo se han utilizado en la práctica.

Se espera que los artículos estén escritos de manera que puedan ser leídos por personas no especialistas en el tema tratado. Se recomienda incluir una lista de lecturas sugeridas para que los lectores no especialistas puedan profundizar en el tema.

Formato del Manuscrito

Los autores deben enviar un archivo en formato PDF del manuscrito que desean someter a referato a:

*Comité Editorial Revista Ingeniería de Sistemas,
Departamento de Ingeniería Industrial,
Universidad de Chile,
Santiago, Chile.
Email: ris@dii.uchile.cl*

Los manuscritos deben estar formateados para hojas tamaño carta, a doble espacio, márgenes de 2,5 centímetros en todos los lados y su extensión no debe exceder las 20 hojas.

La primera hoja debe contener el título del trabajo, nombre y dirección de los autores (teléfono y correo electrónico del autor de contacto) y un resumen de no más de 150 palabras.

Referencias

Las referencias se deben citar en el cuerpo del texto usando el nombre del autor y el año de publicación, e.g., Morton (1998). Al final del artículo se debe incluir la lista en orden alfabético de las referencias citadas en el texto. Para referencias de revistas científicas el formato es el siguiente: Autor(es), Año de publicación. Título. Nombre completo de la revista , Volumen e.g.:

Kodialam, M. y H. Luss, 1998. Algorithms for Separable Nonlinear Resource Allocation Problems. *Operations Research* , 44(2), 272-284.

Para referencias de libros el formato es el siguiente: autor(es), año de publicación. Título. Editorial, Ciudad; e.g.:

Kleinrock, L., 1975. *Queueing Systems* . John Wiley, New York.

En caso de haber más de una referencia con el mismo autor y año de publicación, se debe usar "a", "b", etc. como sufixo del año de publicación para diferenciarlas.

Detalles en www.dii.uchile.cl/ris

PREDICCIÓN DE ARRIBOS Y PROGRAMACIÓN DE TURNOS EN UNA ESTACIÓN DE SERVICIO

OCTAVIO ENEAS BIASOLI * †
JAVIER MARENCO ‡ §

Resumen

En este trabajo se estudian los problemas de predicción de la demanda y de optimización de turnos de trabajo en una estación de servicio, haciendo especial foco en las implicancias de estas decisiones sobre el tiempo de espera de los clientes, la utilización del sistema de atención al público y la cantidad de personal necesario para lograr un buen nivel de servicio. Se ha tomado como objeto de estudio una estación de servicio ubicada en Buenos Aires, Argentina, para comparar su situación actual, realizando horarios de manera manual, con la programación de horarios que se propone en este trabajo. Para ello, se realiza un análisis de los datos históricos de la empresa utilizando herramientas de teoría de colas, se utilizan técnicas de *machine learning* para predecir los arribos de clientes y se plantea un modelo de programación lineal entera para asignar los turnos. Los resultados muestran que las herramientas propuestas en este trabajo permiten mejorar el nivel de servicio reduciendo la dotación de empleados en el caso de estudio analizado.

Palabras Clave: Teoría de colas, Predicción de arribos, Asignación de turnos, Programación lineal entera

*Universidad Torcuato Di Tella, Argentina.

†Universidad Argentina de la Empresa, Argentina.

‡Universidad de Buenos Aires, Argentina.

§Universidad Nacional de General Sarmiento, Argentina.

1. Introducción

La programación de horarios es el proceso de construir un calendario de trabajo para el personal de una organización, de manera tal que la misma pueda realizar sus actividades adecuadamente. A pesar de que es una tarea que influye directamente en los costos y calidad de atención de la empresa, la gestión de horarios rotativos en estaciones de servicio suele ser un tema al que no se le da la importancia que debería tener. En general, la misma es delegada a un encargado de personal que resuelve la problemática de manera manual y sin ningún tipo de análisis. Los horarios normalmente son asignados por intuición y “usos y costumbres” del rubro, sin tener en cuenta la real afluencia de clientes ni la cantidad de personal necesario para atender dicha demanda. De esta manera, es habitual que en algunos días haya más empleados trabajando que los necesarios, que luego faltan en días con mayor concurrencia de clientes. Suele suceder incluso que en algunos días, por algún error humano, falta personal para cubrir los puestos mínimos de trabajo, que hay que completar de urgencia haciendo llamadas de último momento.

Estos problemas de gestión generan costos adicionales que no suelen ser identificados por los empresarios, haciendo que erróneamente releguen el tema en cuestión. Por ello, muchos de ellos se encuentran en este momento pagando sueldos de empleados que no necesitan, perdiendo clientes por brindarles una mala atención y/o padeciendo las consecuencias de un mal clima laboral debido al malestar que la imprevisibilidad genera en los empleados.

Por consiguiente, la pregunta principal de este trabajo consiste en determinar si es posible mejorar los tiempos de espera de los clientes, la utilización del sistema de atención al público y la cantidad de empleados necesarios mediante modelos de optimización combinatoria. Para esto, será necesario predecir la demanda dada por la afluencia de clientes, que constituye unos de los principales insumos para esta decisión.

El presente trabajo está organizado del siguiente modo. En la Sección 2 se realiza una breve revisión bibliográfica de trabajos relacionados. La Sección 3 presenta un análisis de los datos de la empresa aplicando herramientas de teoría de colas para estimar el nivel de atención y los parámetros del sistema. En la Sección 4 se presentan técnicas de *machine learning* para predecir los arribos de clientes y, sobre esta base, en la Sección 5 se introduce un modelo de programación lineal entera para el problema de asignación de turnos. La Sección 6 cierra el trabajo con conclusiones y posibles líneas de trabajo futuro.

2. Estado del arte

La planificación de horarios del personal es un tema que ha sido muy estudiado durante las últimas décadas desde la investigación de operaciones. Uno de los primeros trabajos al respecto es [2], para una empresa que trabaja todos los días del año, planteando como objetivo determinar el mínimo personal necesario en cada día de la semana para satisfacer la demanda. En [3] se estudia el problema de determinar la cantidad de personal necesario en cada turno de cada día de una empresa de servicios australiana que trabaja todos los días del año en horario corrido. En este trabajo, se estimó la demanda futura a partir de los datos de las llamadas telefónicas recibidas durante un año.

Una aplicación habitual de los problemas de asignación de turnos está dada por la planificación del personal en el sistema de salud. Por ejemplo, en [10] se estudia el caso de la asignación de enfermeras en un hospital en Costa de Oro, Australia. En este trabajo aparecen restricciones *soft*, como cantidad máxima y mínima de turnos sin días de descanso, preferencia de turno por parte de cada enfermera, y cantidad máxima y mínima de turnos por semana. En [7] se plantea una situación similar, tomando una cantidad fija de empleados para cada día, cantidad mínima de días de trabajo por empleado por semana, cantidad máxima de días de trabajo corridos por empleado y tiempo mínimo entre turnos como restricciones *hard*. Por otra parte, se considera el agrupamiento de los días de trabajo y el balanceo de la cantidad de descansos en días de semana y fines de semana como restricciones *soft*. Finalmente, en [1] se estudia el caso de unidades de terapia intensiva en un hospital en Estados Unidos, que trabaja todos los días del año en horario corrido. En este caso, el objetivo es reducir los costos de personal a la vez que se balancean los riesgos de asignar empleados de más o de menos a cada turno, recurriendo para ello a un pronóstico de demanda realizado por un modelo de regresión que se nutre de datos históricos del hospital.

Los problemas de optimización combinatorial que surgen en estos trabajos son habitualmente modelados y resueltos por medio de técnicas de programación lineal entera. En los últimos años, además, se hizo frecuente el uso de herramientas de *machine learning* para el análisis y predicción de la demanda, sobre la base de datos históricos. A pesar de que la demanda es incierta, en este trabajo modelaremos el problema como un problema determinístico, suponiendo que la demanda es conocida de antemano. Esto simplifica el modelado y resolución del problema, aunque tiene el costo de que las decisiones

pueden no ser robustas y sufrir las consecuencias de cambios inesperados en el comportamiento de los clientes.

3. Análisis descriptivo

En este trabajo se analiza el caso particular de una estación de servicio consignataria de una marca petrolera internacional, ubicada en Buenos Aires, Argentina, la cual cuenta con seis bocas de despacho de combustible, distribuidas en tres islas. Se cuenta con el detalle de los despachos realizados por el establecimiento durante 38 meses, registrados en forma automática por el sistema informático de la empresa.

La Figura 1 contiene un diagrama de caja con los arribos de clientes en cada hora del día. Desde las 00:00 hasta las 05:00 inclusive, los arribos promedio no superan los 20 autos por hora, los cuales toman una carrera alcista a partir de las 06:00, momento en el cual la demanda crece rápidamente hasta alcanzar un *peak* a las 12:00, con una cantidad de 56 autos por hora en promedio. Luego, la demanda desciende levemente hasta las 15:00 a un promedio de 48 autos por hora, para volver a tomar una tendencia alcista que toca su segundo *peak* del día a las 18:00, con una demanda promedio de 59 autos por hora.

La cantidad de arribos por hora se encuentra entre rangos muy amplios, lo que supone un riesgo para la planificación de horarios de manera manual. Por ejemplo, la diferencia entre el primer cuartil y el tercer cuartil de los arribos a las 19:00 es de 24 autos, lo que haría imposible atender al público correctamente en caso de tener un número demasiado bajo de empleados en ese horario. Asimismo, la diferencia entre el valor mínimo y máximo de este horario es de 117 autos.

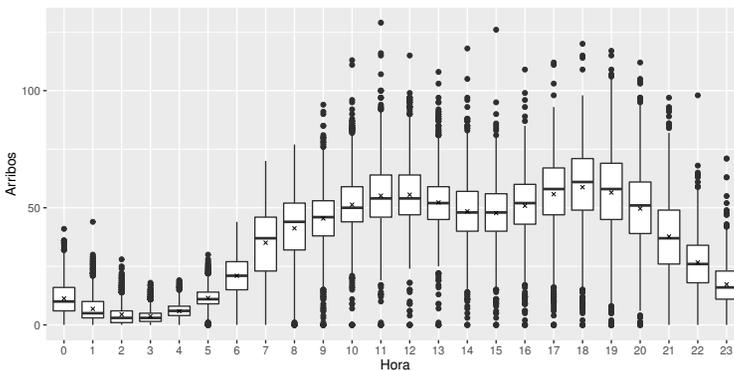


Figura 1: Arribos de clientes por hora.

Dado que en la estación de servicio no se poseen datos sobre el tiempo de servicio, se infirió el mismo a partir de los datos disponibles calculando el tiempo entre despachos en horarios con un flujo continuo de atención. Solamente se tuvieron en cuenta aquellos despachos con un tiempo menor o igual a tres segundos por litro despachado más cinco minutos en concepto de tiempo de cobro. Las mediciones superiores a este valor fueron descartadas por presuntamente incorporar demoras no relacionadas al tiempo de servicio. Se tomaron los datos de despachos de 38 meses, dados por un total de 556067 registros para las seis caras de despacho de la estación de servicio. Este procedimiento arrojó un tiempo de servicio promedio de 3.24 minutos por cliente por cada boca de despacho, con una desviación estándar de 1.50 minutos.

Sin embargo, si se tiene en cuenta que cada empleado puede atender hasta dos caras de un surtidor sin alterar su capacidad de atención, dado que el surtidor puede despachar combustible sin necesidad de supervisión continua por parte de los empleados, se puede considerar que el tiempo de servicio es de la mitad del valor obtenido para una cara. Esta presunción, comúnmente realizada en el rubro, fue validada a través de las cámaras de seguridad del establecimiento, en las cuales se pudo observar que se trata de una suposición razonable. Entonces, suponemos un tiempo de atención de 1.62 minutos por cliente por cada empleado. Con estos datos, tenemos una tasa de servicio de 37.02 clientes por hora por empleado.

Un análisis del proceso de arribos y los tiempos de servicio permite establecer que una cola de tipo $G|G|s$ es un modelo adecuado para este sistema, tomando $s \in \{1, 2, 3\}$ en función de la cantidad de empleados asignados a cada hora. Suponiendo este modelo para el arribo y atención de los clientes, el tiempo de espera en cada hora se aproxima por medio de la fórmula propuesta en [9]. En la Figura 2 se detalla el tiempo promedio de espera en cola en función de la hora del día a lo largo del mes de febrero de 2021. Estos datos corresponden a los resultados obtenidos por la planificación manual de horarios realizada por el establecimiento, que utiliza 8.5 empleados (la empresa tenía en ese momento nueve empleados, pero uno de ellos tomó sus vacaciones anuales durante dos semanas en ese mes).

Este gráfico deja en evidencia dos grandes falencias en cuanto a la planificación manual de los horarios del personal. En primer lugar, se nota que la rigidez de turnos que se utilizan (tres turnos, el primero entre las 06:00 y las 14:00, el segundo entre las 14:00 y las 22:00, y el tercero entre las 22:00 y las 06:00) carece de la flexibilidad adecuada para captar los distintos *peaks* de demanda a lo largo del día. Específicamente, las 11:00, 12:00 y 18:00 son momentos del día que necesitarían especial atención para evitar que el tiempo de espera aumente en forma inaceptable.

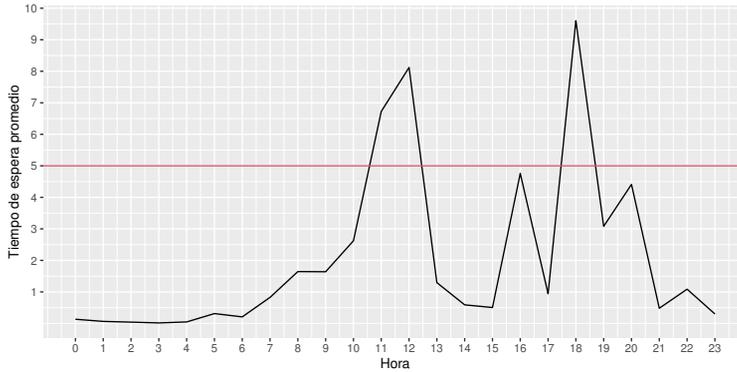


Figura 2: Tiempo promedio de espera por hora con la planificación manual.

En segundo lugar, más allá de la rigidez de los turnos utilizados, sale a la luz que hay un problema de asignación de personal, ya que en los horarios mencionados anteriormente debería haber un empleado adicional para evitar que los tiempos de espera en cola se disparen. Habiendo realizado un análisis de la capacidad de atención de la empresa, puede afirmarse que el cuello de botella es totalmente atribuible a este problema de planificación de horarios más que a una falta de la capacidad instalada del establecimiento. Esto se debe a que la capacidad máxima de la empresa, que consiste en tener trabajando a tres empleados (servidores) de manera simultánea, permite atender el 99.95 % de las demandas históricas, o el 99.32 % de las demandas históricas si se establece como tope una utilización del 80 % de la capacidad máxima. Este valor de referencia es muy importante para este análisis, ya que una vez superado el umbral de 80 % de utilización del sistema, el impacto sobre el tiempo de espera es muy significativo.

En el 4.46 % de las 672 horas que componen el mes de febrero de 2021 el tiempo de espera fue superior a cinco minutos, lo que significa que en una hora cada 22 horas la estación de servicio brindó una calidad de atención al público que se encuentra por debajo de los estándares de la industria. Por otro lado, el tiempo de espera promedio a lo largo del mes fue de 1.91 minutos.

El nivel de utilización promedio del sistema es de 43.24 %. Sin embargo, el 6.85 % de los horarios del mes se encuentran en un nivel de utilización superior al 80 %, de los cuales el 10.87 % se encuentra por encima del 100 %. Evidentemente, es importante para la empresa mejorar estos indicadores, dado que existe una pérdida económica asociada a un bajo nivel de servicio, y esta observación es la que motiva el presente trabajo.

4. Predicción de arribos

Con la información proporcionada por la empresa se creó un *dataset* con 27.624 observaciones, 16 variables predictoras (descritas en la Tabla 1) y una variable de respuesta (cantidad de arribos por hora).

Variable	Descripción	Valores
Hora	Hora del día	$\{0, \dots, 23\}$
Día	Día del mes	$\{1, \dots, 31\}$
Mes	Mes del año	$\{1, \dots, 12\}$
Año	Número de año	$\{2018, \dots, 2021\}$
Día de la semana	Día de la semana	$\{\text{Lun}, \dots, \text{Dom}\}$
Número de semana	Número de semana en el año	$\{1, \dots, 52\}$
Camión	Especifica si estaba presente el camión de descarga de combustible	$\{\text{Sí}, \text{No}\}$
Fin de semana	Especifica si el día es sábado o domingo	$\{\text{Sí}, \text{No}\}$
Feriado	Especifica si el día es feriado	$\{\text{Sí}, \text{No}\}$
Estación	Estación del año	$\{\text{Ver}, \dots, \text{Prim}\}$
C. Precio	Especifica si es el día anterior a un aumento de precios	$\{\text{Sí}, \text{No}\}$
Cerrado	Especifica si la estación de servicio está cerrada en ese momento	$\{\text{Sí}, \text{No}\}$
Pandemia	Especifica el nivel de restricciones de movilidad dadas por la cuarentena de 2020, desde 0 (sin restricciones) hasta 5 (máximo de restricciones)	$\{0, \dots, 5\}$
Temperatura	Temperatura promedio de esa hora	\mathbb{R}
Precipitaciones	Registro de lluvias (en mm.) en esa hora	\mathbb{R}_+
Tendencia	Días transcurridos desde el 1/1/2018	\mathbb{Z}_+

Tabla 1: Variables del *dataset*.

Al momento de analizar los atributos significativos para predecir la cantidad de arribos por hora en una estación de servicio, el día de la semana es un dato crucial. La Figura 3 muestra que, en promedio, los arribos por hora de los lunes, martes, miércoles y jueves son muy similares, pero con una pequeña tendencia alcista del primero al último. En vísperas del fin de semana se produce un salto importante en el movimiento del establecimiento, que llega a su punto máximo los sábados. Los domingos representan el día de la semana con menor tráfico, siendo apenas superior a 30 autos por hora.

Así como el día de la semana tiene un efecto sobre los hábitos de consumo

de los clientes de la estación de servicio, se buscó determinar si cada uno de los días que componen al mes también lo tiene. En general, los dueños de estaciones de servicio suelen considerar que la primera quincena del mes tiene mayor afluencia de clientes que la segunda quincena. Sin embargo, los datos muestran que la cantidad de arribos promedio por hora a lo largo del mes tiene una gran variación, y no presenta un patrón tan marcado. La Figura 4 reporta estos datos. Puede verse que la primera quincena del mes suele tener una cantidad de arribos más estable, comenzando con su menor valor promedio el día 1 y alcanzando su *peak* el día 7. Por el otro lado, la segunda quincena tiene mayor dispersión, teniendo su menor valor promedio el día 25 y su *peak* el día 31, valores que, a la vez, son el piso y techo del mes.

Otro de los atributos que demostró tener un impacto importante en la predicción de arribos es la presencia o no del camión cisterna, que es aquel que se presenta en la estación de servicio para realizar la descarga de combustible y llenar los tanques de almacenamiento. Este atributo pudo ser incorporado gracias a que la estación de servicio cuenta con un registro de la hora de llegada y de la hora de salida del camión cisterna.

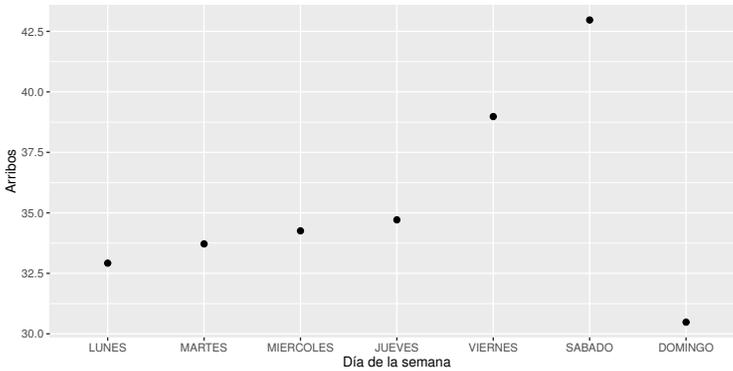


Figura 3: Cantidad de arribos en función del día de la semana.

Finalmente, a partir de datos obtenidos del Servicio Meteorológico Nacional, en cuanto a temperatura y precipitaciones de cada hora del período bajo análisis, se pudo identificar la relativa importancia que tiene la temperatura para predecir los arribos por hora, dado que registró una correlación positiva entre la temperatura a cada hora y la cantidad de arribos que suceden en la misma.

Para llevar a cabo la predicción de arribos el *dataset* se dividió en tres partes. Por un lado, se creó el conjunto de testeo, compuesto por la totalidad de los datos desde 01/02/2021 hasta 28/02/2021 inclusive, que fue separado y utilizado únicamente para verificar la performance final del modelo cuando

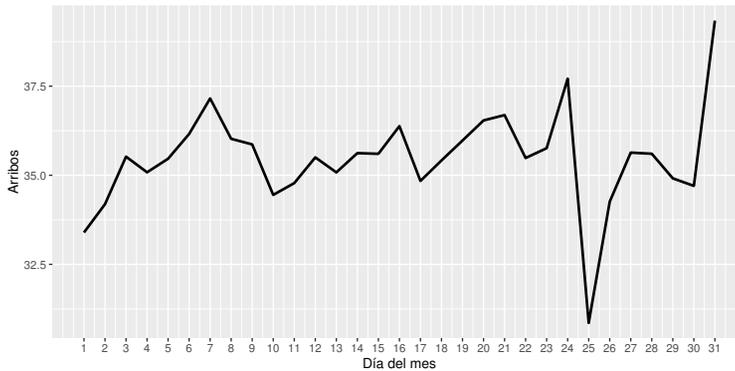


Figura 4: Cantidad de arribos en función del día del mes.

se consideró que el mismo ya se había entrenado lo suficiente. Estos datos se asumen desconocidos a los fines de este trabajo, ya que la predicción de los arribos durante este período de tiempo se utilizará como *input* para los experimentos reportados en la próxima sección.

Por otro lado, se creó un conjunto de validación, compuesto por el 25 % de los datos entre 01/02/2020 y 31/01/2021 inclusive, seleccionados de manera aleatoria. De esta manera, se busca tener en cuenta la temporalidad de los datos, intentando que el modelo sea preciso a la hora de predecir aquellos más recientes en el tiempo, especialmente teniendo en cuenta los efectos de la cuarentena impuesta durante 2020. Al cambiar de alguna manera los hábitos de consumo de los clientes como, por ejemplo, la cantidad de clientes que utilizan sus vehículos y que posteriormente necesitan cargar combustible, se consideró conveniente diseñar el conjunto de validación con estas características. Este conjunto de validación simuló ser un grupo de datos desconocidos durante la etapa de entrenamiento, que sirvió para ir confirmando cada una de las decisiones que se tomaron en cuanto a ingeniería de atributos y selección de modelos.

Por último, se creó un conjunto de entrenamiento, compuesto por los datos desde 01/01/2018 hasta 31/01/2021 inclusive que no fueron seleccionados al azar para formar parte del conjunto de validación.

Para realizar esta estimación de la manera más precisa posible se utilizaron dos algoritmos de distintas características, un algoritmo de tipo *random forest* [4] y un algoritmo de *gradient boosting machine* [5]. Luego de optimizar los hiperparámetros de ambos métodos, se arribó a los siguientes resultados. Con la muestra de entrenamiento, el algoritmo *random forest* obtuvo un error cuadrático medio de 60.42 sobre los datos de validación o, lo que es equivalente, un error promedio de 7.77 autos en su predicción correspondiente a cada

hora. Por otro lado, con la muestra de entrenamiento, *gradient boosting machine* demostró tener un error cuadrático medio de 50.04 sobre los datos de validación, o de 7.03 autos por hora.

Dado que *gradient boosting machine* demostró tener un desempeño un 17% superior al que obtuvo *random forest*, se seleccionó este algoritmo para realizar la predicción de arribos de períodos futuros. El error cuadrático medio de *gradient boosting machine* sobre los datos de testeo, que son los que corresponden al mes de febrero de 2021, fue de 57.46, o de 7.58 autos por hora. La Figura 5 presenta un gráfico de dispersión entre las predicciones de *gradient boosting machine* y las observaciones. En esta figura, la línea roja corresponde a la recta $y = x$, de modo tal que las predicciones son más precisas cuanto más cercanas se encuentren a la línea roja. Se puede observar un ajuste aceptable de las predicciones del algoritmo con relación a los datos reales. Posteriormente, estos datos serán utilizados como entrada del modelo de planificación de horarios que se presenta en la próxima sección.

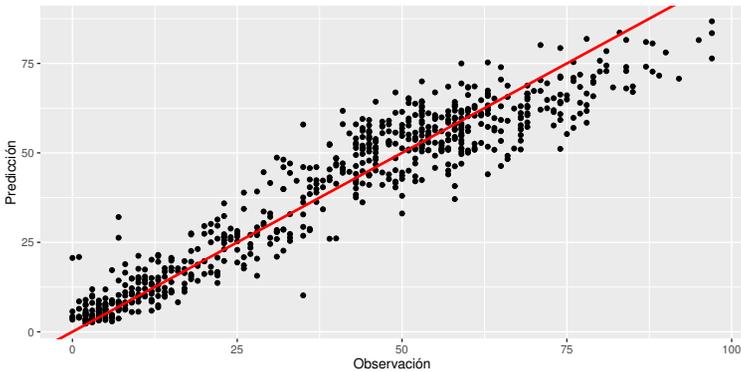


Figura 5: Comparación entre las predicciones de *gradient boosting machine* y las observaciones de arribos.

Es interesante mencionar que al momento de predecir la cantidad de autos que van a ingresar a la estación a cada hora, no todas las variables tienen la misma entidad. Algunas variables predictoras son más relevantes a la hora de disminuir la tasa de error lo máximo posible, y esto es algo que *gradient boosting machine* permite conocer. Los atributos más importantes para la estimación son la hora del día, el día de la semana, el nivel de cuarentena (para datos de 2020 y 2021) y el número de semana del año, que representan el 90.98% de la influencia relativa de las variables.

5. Optimización de turnos

Se describe en esta sección un modelo de programación lineal entera para el problema de asignar turnos a los empleados durante un mes completo, con el objetivo de cubrir la demanda de la mejor manera posible y además tener una asignación que sea aceptable para los empleados.

Se cuenta con un conjunto de empleados, que realizan turnos de 8 horas. No es necesario recurrir a todos los empleados para realizar la asignación, dado que uno de los objetivos es minimizar la cantidad total de empleados contratados. Cada empleado puede comenzar su turno en cualquier hora entre las 6:00 y las 16:00, y entre las 22:00 y las 0:00. Estas restricciones están dadas por el requerimiento de que ningún empleado comience o termine su turno entre la 1:00 y las 5:00, de modo tal que los empleados puedan conseguir con mayor facilidad transporte público, en caso de necesitarlo, y de tener una mejor experiencia laboral. Es deseable pero no obligatorio que los turnos de cada empleado se mantengan razonablemente constantes a lo largo de los días. Finalmente, cada empleado tiene un día de descanso cada 7 días.

Para el modelo tenemos los siguientes conjuntos:

- $E = \{1, \dots, n\}$: conjunto de empleados,
- $H = \{0, \dots, m\}$: horas del mes, representadas en forma consecutiva desde las 0 hs del primer día del mes (por ejemplo, para un mes de 30 días se tiene $m = 719$),
- $H_0 \subseteq H$: horas del mes en las que no se puede comenzar un turno, y que corresponden a ingresos o egresos de turnos entre la 1:00 y las 5:00.

Definimos el parámetro $a_h \in \mathbb{R}$ para cada hora $h \in H$, que representa la cantidad de arribos de clientes esperada en la hora h del intervalo de planificación. Definimos también el parámetro $b_e \in \mathbb{Z}_+$ para cada empleado $e \in E$, que especifica el horario de ingreso del empleado en el último día del mes anterior o bien $b_e = 0$ si el empleado no trabajó ese día. Finalmente, definimos el parámetro $c_e \in \mathbb{Z}_+$ para cada empleado $e \in E$, que especifica cuántos días lleva trabajados de manera continua al final del mes que se cierra, de modo tal que el empleado e debe tener al menos un día de descanso entre los primeros $d_e := \max\{1, 7 - c_e\}$ días del mes inmediato siguiente.

Para plantear el modelo, introducimos las siguientes variables binarias:

- x_e : especifica si el empleado e será contratado, para $e \in E$,

- y_{eh} : especifica si el empleado e comienza un turno en la hora h , para $e \in E$ y $h \in H$,
- z_{eh} : especifica si el empleado e comienza un turno en la hora h y si en el día anterior también comenzó un turno a la misma hora del día, para $e \in E$ y $h \in \{24, \dots, m\}$.

Con estas definiciones, podemos plantear el siguiente modelo:

$$\text{mín} \quad \sum_{e \in E} x_e - \varepsilon \sum_{e \in E} \sum_{h=24}^m z_{eh} \quad (1)$$

$$\sum_{h \in H} y_{eh} = 24 x_e \quad \forall e \in E \quad (2)$$

$$\sum_{h'=h-19}^h y_{eh'} \leq 1 \quad \forall e \in E, \forall h \in H, h \geq 19 \quad (3)$$

$$y_{eh} = 0 \quad \forall e \in E, \forall h \in H, h \leq b_e - 5 \quad (4)$$

$$\sum_{h'=h-167}^h y_{eh'} \leq 6 \quad \forall e \in E, \forall h \in H, h \geq 167 \quad (5)$$

$$\sum_{h=0}^{24 d_e - 1} y_{eh} \leq d_e - 1 \quad \forall e \in E \quad (6)$$

$$y_{eh} = 0 \quad \forall e \in E, \forall h \in H_0 \quad (7)$$

$$\tau \sum_{e \in E} \sum_{h'=h-7}^h y_{eh'} \geq a_h \quad \forall h \in H, h \geq 7 \quad (8)$$

$$z_{eh} \leq y_{e, h-24} \quad \forall e \in E, \forall h \in H, h \geq 24 \quad (9)$$

$$z_{eh} \leq y_{eh} \quad \forall e \in E, \forall h \in H, h \geq 24 \quad (10)$$

$$x_e \in \{0, 1\} \quad \forall e \in E \quad (11)$$

$$y_{eh} \in \{0, 1\} \quad \forall e \in E, \forall h \in H \quad (12)$$

$$z_{eh} \in \{0, 1\} \quad \forall e \in E, \forall h \in H, h \geq 24 \quad (13)$$

La función objetivo (1) busca minimizar la cantidad de empleados necesarios para atender al público a la vez que se mantiene el mismo horario de inicio de la mayor cantidad posible de turnos respecto del día anterior. En este contexto, ε es una constante suficientemente pequeña para priorizar el primer término de la función objetivo. Las restricciones (2) especifican que los empleados que sean contratados deben trabajar exactamente 24 días en el mes. Las restricciones (3) piden que los empleados tengan al menos 12 horas de descanso entre el final de una jornada laboral y el inicio de la siguiente, mientras que las restricciones (4) aseguran este tiempo de descanso entre el último día del mes que finaliza

y el primer día del mes que comienza. Las restricciones (5) solicitan que no se asignen más de seis días de trabajo seguidos sin descanso durante el mes, mientras que las restricciones (6) solicitan este día de descanso en función de los días trabajados al final del mes anterior. Las restricciones (7) piden que no se comiencen turnos en los horarios prohibidos. Las restricciones (8) especifican que la capacidad de atención al público de la empresa, de 37.02 autos por empleado por hora, no puede superar el 80% de utilización, y entonces se toma $\tau = 0,8 \times 37,02$. Las restricciones (9)-(10) ligan las variables z con las variables y , y finalmente las restricciones (11)-(13) especifican la naturaleza de las variables.

Se implementó el modelo en el lenguaje de modelado ZIMPL [8] y se resolvió con el solver SCIP [6] con un tiempo máximo de ejecución de 5 minutos. Luego de este tiempo, la solución obtenida tiene un gap de optimalidad de 0.65%, lo cual se considera aceptable para los objetivos de este trabajo. La Figura 6 muestra la solución para el mes de febrero, especificando los horarios de inicio de cada empleado (las celdas marcadas con “F” corresponden a días de descanso, llamados *francos* en este contexto). Para esta ejecución se asume $b_e = c_e = 0$ para todo $e \in E$, es decir que comenzamos una planificación “desde cero”. Es interesante observar que los horarios de ingreso son relativamente estables, confirmando que el objetivo planteado es razonable para este problema. Este atributo es muy valorado por el personal que trabaja bajo jornadas rotativas, lo que permite que se cree un mejor clima laboral en la empresa. Además, es importante destacar que el modelo ha arribado a una solución con sólo ocho empleados, lo que implica una reducción del 6% en los costos de personal con respecto a la planificación manual.

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	
Día	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Empleado 1	16	16	16	16	22	22	F	16	16	16	16	22	22	F	16	16	16	16	22	22	22	F	16	16	16	16	22	22	22
Empleado 2	8	8	F	8	8	8	8	8	8	F	8	8	8	8	8	8	F	8	8	8	8	8	8	8	F	9	9	9	9
Empleado 3	F	16	16	16	16	16	16	F	16	16	16	16	16	16	F	16	16	16	16	16	16	16	F	16	16	16	16	16	16
Empleado 4	0	0	0	0	0	0	F	0	0	0	0	0	0	F	0	0	0	0	0	0	0	F	0	0	0	0	0	0	F
Empleado 5	0	F	9	9	16	16	16	16	F	9	9	16	16	16	16	F	9	9	16	16	16	16	16	16	F	9	16	16	16
Empleado 6	14	14	14	14	14	14	F	14	14	14	14	14	14	F	14	14	14	14	14	14	14	F	14	14	14	14	14	14	F
Empleado 7	6	6	6	F	8	8	8	6	6	6	F	8	8	8	6	6	6	F	8	8	8	6	F	8	8	8	8	8	8
Empleado 8	6	F	6	6	6	6	6	6	F	6	6	6	6	6	F	6	6	6	6	6	6	6	F	6	6	6	6	6	6

Figura 6: Planificación para el mes de febrero.

Evaluamos el tiempo de espera en cola de los clientes con la asignación de turnos proporcionada por el modelo de la misma forma que se realizó para la planificación manual, suponiendo un modelo $G|G|s$ y utilizando la aproximación del tiempo de espera descrita en [9]. Se muestran estos resultados en la

Figura 7. El tiempo de espera en cola a partir de la asignación de personal que plantea el modelo otorga resultados interesantes a lo largo del día, oscilando en valores alrededor de un solo minuto de espera, o menos, en promedio a cada hora. El único momento del día en que el tiempo de espera en cola es significativamente superior al resto de los horarios es a las 18:00, siendo el mismo de aproximadamente 4.50 minutos.

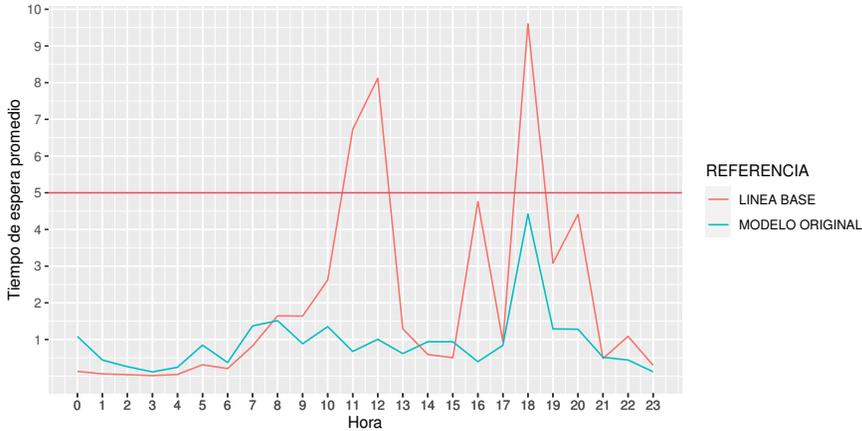


Figura 7: Tiempo promedio de espera por hora con la planificación manual (serie en rojo) y con la planificación obtenida por el modelo (serie en azul).

Dado que la predicción realizada por *gradient boosting machine* no es exacta, hay unas pocas horas en el mes en las cuales la utilización supera el 80%, generando que los tiempos de espera se acerquen a 11 minutos. Sin embargo, tal como se concluyó anteriormente, esto no afecta significativamente la performance global del modelo a lo largo del mes bajo análisis, ya que todos los tiempos de espera promedio a cada hora se encuentran por debajo de los cinco minutos.

El nivel de utilización promedio del sistema es de 43.41%, prácticamente idéntico a lo observado en la estación de servicio con su planificación manual. Sin embargo, solo en algunos pocos momentos del mes el nivel de utilización es superior al 80% con esta metodología, fracción que representa el 3.13% del total, de los cuales el 4.80% se encuentra por encima del 100%. En otras palabras, solo en una oportunidad la capacidad de atención de la empresa fue superada por la demanda del momento.

Se experimentó con un segundo escenario, en el que el nivel máximo de utilización del sistema se fija en el 76% en lugar del 80%. Esto proporciona un mejor nivel de servicio, pero potencialmente estresando la asignación de turnos. En este caso, el solver termina con un gap de optimalidad de 0.47% luego de 5 minutos de ejecución. El tiempo de espera promedio es de 0.76

minutos a lo largo del mes bajo análisis, obteniendo una ventaja considerable con respecto al escenario original, que consiguió atender a los clientes en 0.92 minutos en promedio. Asimismo, este nuevo escenario presenta tiempos de espera superiores a cinco minutos en solo el 0.74 % de los casos, a diferencia del escenario original, que llega a 1.93 %. Esta diferencia de rendimiento se origina por la holgura adicional en la capacidad que el escenario ajustado tiene por sobre el escenario original.

Si, por el contrario, se permite una utilización máxima del sistema del 90 %, la solución obtenida por el modelo utiliza un empleado menos que los dos escenarios anteriores, pero el tiempo de espera promedio de los clientes asciende a 2.86 minutos, un valor que triplica los 0.92 minutos del escenario original. Asimismo, a lo largo del mes, en este escenario relajado se supera la barrera de los cinco minutos de espera en el 6.25 % de los casos.

6. Conclusiones y trabajo futuro

El personal es el concepto de mayor peso dentro de la estructura de costos de una estación de servicio, por lo que su correcta gestión es indispensable para asegurar la rentabilidad del negocio. Asimismo, es el personal también quien, a través de su atención, conquista y retiene a los clientes de la empresa. De esta manera, surgen dos objetivos contrapuestos que ponen a prueba a los empresarios del sector: cómo maximizar la calidad de atención manteniendo bajo control los costos de personal.

El modelo planteado atiende estos dos objetivos minimizando la cantidad de empleados (y en segundo lugar la variación de sus turnos), y manteniendo un nivel de utilización máximo como restricción. En el ejemplo analizado, el modelo permite concluir que la demanda del mes estudiado se puede atender con ocho empleados, proporcionando al mismo tiempo un nivel de servicio aceptable. Este dato es fundamental, ya que orienta al empresario en cuanto a la inversión necesaria para atender correctamente a sus clientes y le garantiza que la misma ha sido calculada de manera profesional.

Estos resultados fueron posibles en virtud del análisis de los datos existentes en la empresa. El sistema informático de despacho registra esta información, que resulta crucial para la asignación de turnos.

Sería interesante estudiar posibles refinamientos al modelo para tener en cuenta otros aspectos que puedan contribuir al buen clima laboral, como por ejemplo realizar una asignación equitativa de los turnos menos preferidos por los empleados, o bien tener en cuenta las preferencias de cada empleado en la

asignación. Sería relevante también estudiar mecanismos para reforzar el modelo, dado que no fue posible resolverlo en forma óptima en el límite de tiempo especificado. Posiblemente el uso de un solver más potente permita subsanar esta situación, aunque también sería interesante lograr soluciones óptimas con el solver utilizado en este trabajo.

Agradecimientos. Los autores quisieran agradecer al revisor anónimo por sus interesantes y constructivos comentarios, que ayudaron a mejorar la versión inicial de este trabajo.

Referencias

- [1] O. T. Aydas, A. D. Ross, M. C. Scanlon, y B. Aydas. New results on integrated nurse staffing and scheduling: The medium-term context for intensive care units. *Journal of the Operational Research Society*, 72(12):2631–2648, 2021.
- [2] K. R. Baker. Scheduling full-time and part-time staff to meet cyclic requirements. *Journal of the Operational Research Society*, 25(1):65–76, 1974.
- [3] N. Beaumont. Scheduling staff using mixed integer programming. *European journal of operational research*, 98(3):473–484, 1997.
- [4] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [5] J. H. Friedman. Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4):367–378, 2002.
- [6] G. Gamrath, D. Anderson, K. Bestuzheva, W.-K. Chen, L. Eifler, M. Gasse, P. Gemander, A. Gleixner, L. Gottwald, K. Halbig, et al. The scip optimization suite 7.0. 2020.
- [7] M. S. Gondane y D. Zanwar. Staff scheduling in health care systems. *IOSR Journal of Mechanical and Civil Engineering*, 1(6):28–40, 2012.
- [8] T. Koch. Rapid mathematical programming. 2005.
- [9] H. Sakasegawa. An approximation formula $l = ap^q / (1-p)$. *Ann. Inst. Statist. Math., Part A*, 29:67–75, 1977.
- [10] J. Thornton y A. Sattar. Nurse rostering and integer programming revisited. En *International conference on computational intelligence and multimedia applications*, páginas 49–58. Citeseer, 1997.

INFERENCIA BAYESIANA DE UN MODELO MARKOVIANO DE FÚTBOL CON APLICACIÓN EN SCOUTING

PABLO GALAZ*
SEBASTIÁN MENA*
DENIS SAURÉ*

Resumen

En este trabajo proponemos un enfoque analítico que utiliza datos granulares de fútbol profesional para modelar encuentros de fútbol considerando interacciones a nivel de jugador para predecir el desarrollo de un encuentro. Este enfoque de modelamiento representa un avance significativo respecto a la literatura, enfocada hasta ahora principalmente en predecir de manera agregada el resultado de un encuentro, puesto que permite analizar la influencia de jugadores individuales en el rendimiento colectivo de un equipo, por lo que cuenta con múltiples aplicaciones en la administración de un equipo de fútbol, como por ejemplo el proceso de scouting. El modelo propuesto visualiza el desarrollo de un partido como una cadena de Markov en tiempo discreto, en el cual las probabilidades de transiciones dependen de una forma no trivial en parámetros que hablan de las características cognitivo-perceptuales y técnicas de los jugadores. El enfoque propuesto utiliza inferencia bayesiana para estimar la distribución a posterior de los parámetros que definen las características de los jugadores. Ilustramos la factibilidad y el uso potencial de este enfoque utilizando datos de la temporada 2017-2018 de la Premier League de Inglaterra. Una vez calibrado, utilizamos el modelo propuesto para realizar múltiples análisis de sensibilidad, propios de las labores asociadas al proceso de scouting. Nuestros resultados hablan del gran potencial para la práctica del uso este tipo de modelos en particular, y de *sports analytics* en general.

Palabras Clave: Inferencia bayesiana, Modelo markoviano, Simulación, Scouting, Fútbol.

*Departamento de Ingeniería Industrial, Universidad de Chile

1. Motivación

La disciplina de *sports analytics* se centra en el uso de datos y estadísticas para apoyar la toma de decisiones en el deporte mediante modelos analíticos. Si bien su uso es un lugar común desde ya hace más de una década en las ligas profesionales más importantes de Estados Unidos, su uso en el fútbol se vio un tanto demorado precisamente por la falta de datos detallados acerca del transcurso de un evento deportivo. Sin embargo, recientemente han comenzado a surgir distintas técnicas para registrar los *eventos*, jugada-a-jugada, que ocurren durante un partido de fútbol, lo que permite analizar el rendimiento individual de los futbolistas. Actualmente, el nivel más detallado de datos, conocido como *eventing*, consiste en un registro de cada movimiento del balón durante un partido, incluyendo los jugadores involucrados, el minuto de juego, el lugar del campo, en conjunto con un número de métricas de desempeño, como por ejemplo, el *expected goal*, asociado a los tiros a portería [8]. El uso de técnicas es cada vez más masivo, al punto que actualmente es posible acceder a datos con este nivel de detalle para la mayoría de las ligas profesionales de fútbol del mundo. Entre los actuales proveedores de este tipo de data a nivel mundial se encuentran OPTA, Instat, WyScout, por mencionar algunos.

La abundancia de los datos, junto al gran grado de detalle de estos, ha propiciado un desafío de *Big Data* en el tratamiento de esta información, el que es posible enfrentar utilizando herramientas analíticas. A modo de ejemplo, a nivel táctico, unas de las decisiones más importantes en fútbol son aquellas acerca de la disposición espacial de los jugadores durante diferentes etapas del juego [19]. En este contexto, utilizando herramientas analíticas es ahora posible determinar y analizar las zonas de influencias de los futbolistas en fase defensiva, considerando su dirección y velocidad de desplazamiento, lo que permite identificar zonas de pases con diferentes niveles de riesgos y, así, poder explotar zonas de bajo control del rival. De forma similar, a nivel estratégico, es posible utilizar la información disponible para categorizar a los futbolistas de acuerdo a sus perfiles y a su desempeño, lo que posibilita identificar perfiles de futbolistas atractivos para un club, un input crucial para realizar el proceso de búsqueda de nuevos jugadores, conocido como *scouting* [20].

Tradicionalmente, las unidades técnicas de análisis en los clubes de fútbol dependen fuertemente del análisis de vídeo para realizar labores de scouting, en un enfoque más bien cualitativo de evaluación y búsqueda de futbolistas. Sin embargo, los entrenadores y analistas típicamente son capaces de recordar

menos de la mitad de los incidentes claves que surgen durante un partido [11]. Este recuerdo se ve afectado principalmente por varios factores: en general se tiende a seguir el balón y se pierde información lejana a él; puntos de vistas y prejuicios establecidos, pues algunos entrenadores solo ven lo que quieren o esperan ver; efectos de las emociones, como el estrés y la ira, que afectan a la concentración y pueden distorsionar la impresión del entrenador sobre el partido; además de las limitaciones propias de la memoria humana. El análisis cuantitativo de los datos surge entonces como una forma de evidenciar o evitar sesgos cualitativos mencionados. Para ello es necesario postular modelos analíticos que traduzcan datos en información útil. La irrupción de este tipo de modelos causó la revolución de sport analytics en el béisbol, documentado por Michael Lewis en su libro *Moneyball* [10], donde cuenta como General Manager de los Oakland Athletics adoptan una política de contratación basada en datos obteniendo rendimientos por sobre lo esperado. En el básquetbol tenemos el ejemplo del proyecto *Carmelo* desarrollado por *FiveThirtyEight*, el cual busca predecir el rendimiento futuro de los deportistas en base a modelos probabilísticos con datos históricos de la NBA [5].

1.1. Objetivo y supuestos.

En este trabajo realizamos una primera aproximación al desarrollo de una herramienta, basada en un modelo analítico, que permita apoyar el proceso de *scouting* de los clubes de fútbol. Para eso, mediremos el impacto que genera la inclusión de un futbolista en un club, en reemplazo de otro, en el rendimiento deportivo del equipo, en base a métricas agregadas del rendimiento del equipo. Una de las métricas utilizadas, de gran interés para la dirigencia del club, es la probabilidad de ser campeón de un torneo. Un punto de partida en nuestro modelo es reconocer la heterogeneidad de los jugadores en términos tanto de sus perfiles técnicos como de sus procesos de toma de decisión dentro del campo de juego. Esta última característica, de aspecto cognitivo-perceptual, permite capturar la diferencia en desempeño entre jugadores de similares capacidades técnicas. Esto es, modelamos la toma de decisiones de los futbolistas condicionado a distintas circunstancias propias del juego y, en el ámbito técnico, modelamos la capacidad de ejecución de la acción decidida por el jugador.

En nuestro modelo, utilizamos las características de los jugadores para trazar, de forma estilizada/simplificada, el transcurso de un partido de fútbol. Para esto, modelamos un encuentro como un proceso Markoviano, donde las características cognitivas-perceptuales y técnicas de cada futbolista determinan, de manera probabilista, las transiciones entre distintos estados de fase del

juego. En particular, en términos de toma de decisiones durante un partido, los jugadores constantemente deben escoger (de manera probabilista) entre tres alternativas: intentar un pase, intentar un regate o rematar al arco. En términos de la ejecución técnica, la capacidad técnica del jugador y del rival que se enfrenta en el momento (ya sea arquero, defensa rival, etc.) determina (de forma probabilista) el éxito o no de la acción.

Nuestro modelo base de un partido de fútbol depende de múltiples parámetros, que gobiernan las acciones de los jugadores y sus resultados. Estos parámetros deben ser inferidos desde la data, para poder utilizarse en el análisis. Para esto utilizaremos el paradigma de inferencia Bayesiana: asumiremos que los parámetros desconocidos son aleatorios, y los equiparemos con distribuciones a priori; luego calcularemos la distribución a posterior de los parámetros, condicionales en la data, y utilizaremos dicha distribución para estimar los parámetros, por ejemplo, tomando sus valores esperados. Para modelar la toma de decisión de los jugadores y el éxito de dichas acciones utilizamos modelos de decisión logística multinomial. Esto resulta en modelos donde la inferencia no se puede realizar en forma analítica, por lo que utilizamos la técnica numérica conocida como Monte Carlo Markov Chain (MCMC) para aproximar las distribuciones a posterior de los parámetros en base a simulación: ver, por ejemplo [13].

Ilustramos la aplicación de nuestro modelo utilizando datos de nivel *eventing* correspondientes a todos los partidos de la temporada 2017-2018 de la *Premier League* (primera división de Inglaterra) [16]. Recordamos que éstos datos mantienen un registro de cada interacción ocurrida durante el partido que involucre al balón, incluido en que coordenadas del campo de juego ocurre el evento, el tiempo de juego, el jugador asociado, si la acción fue exitosa o no y, en caso de aplicar, jugadores rivales asociados.

1.2. Contribución y resultados.

La principal contribución hecha por nuestro trabajo consiste en el desarrollo de una herramienta que permite medir el impacto absoluto y/o relativo de incluir un jugador en un equipo determinado, en el rendimiento global de un equipo en una competencia. Esto tiene un potencial significativo para mejorar inmensamente los procesos de scouting en el fútbol profesional en países como Chile, y en toda Sudamérica. Es importante notar que, por lo menos en la literatura académica (más detalles en la siguiente sección), el tipo de modelos de transcurso de partidos desarrollados en el pasado pone el foco en la predicción del resultado global del encuentro; dichos modelos no son de manera directa capaces de responder preguntas fundamentales relacionadas a scouting

y a caracterizar la habilidad de los jugadores en términos multidimensionales.

Una segunda contribución, quizás tanto o más importante, aparece como subproducto del proceso de calibración de nuestro modelo: la herramienta permite caracterizar individualmente los perfiles cognitivos-perceptuales y técnicos de cada jugador, posibilitando un análisis detallado y multidimensional de los jugadores. Esto permite, por un lado, ampliar el espectro de decisiones que pueden ser apoyadas por análisis de este estilo, y por otro, el conocer mejor a los jugadores de un plantel, realizar reforzamientos de algunas habilidades técnicas, o focalizar el estilo de juego individual en aspectos donde los jugadores cuentan con ventajas comparativas.

En relación con lo anterior, y a modo de ejemplo, nuestro modelo permite identificar el efecto de la inclusión de un futbolista en un equipo en el cambio de la probabilidad de ser campeón: incluir a Mohamed Salah en el Stoke City, aumenta la probabilidad de ser campeón de ese equipo en un 6.34%; la inclusión de Kevin de Bruyne en el Arsenal, aumenta su probabilidad de ser campeón en un 2.46%. El modelamiento innovador y detallado del transcurso de un partido en función de las habilidades individuales de los jugadores identifica, por ejemplo, qué jugador es más beneficioso para cada uno de los clubes en base a sus propias necesidades y estilos de juego, lo que se traduce en un scouting basado en criterios objetivos y cuantificables. También, abre la posibilidad de medir el valor real de un futbolistas en términos monetarios para un club, en base a los objetivos y potenciales logros que pueda obtener, no solo salir campeón, sino que clasificar a torneos internacionales o evitar el descenso.

1.3. Estructura del manuscrito.

En la Sección 1.3 repasaremos la literatura relacionada y en la Sección 1.3 presentamos el modelo propuesto, junto al esquema de inferencia Bayesiana. En la Sección 3.3 mostraremos los resultados de calibración, e ilustramos el análisis posible de realizar utilizando nuestra propuesta. Finalmente, en la sección 4.2, presentamos una discusión de nuestros supuestos y resultados, junto a nuestras conclusiones.

2. Revisión de la Literatura

La estadística ha ido de la mano del fútbol desde su comienzo, principalmente con el desarrollo de las ligas profesionales. Sin embargo, es tan solo en los 80's que comenzaron a surgir los primeros esfuerzos para registrar, mas allá de es-

tadísticas agregadas, los sucesos que iban ocurriendo en los partidos de fútbol, con el objetivo de analizar a un potencial equipo rival o analizar el rendimiento individual de un futbolista [4, 9, 17]. La tecnología asociada a estos esfuerzos ha evolucionado desde registro manual, pasando por grabaciones de sonido y el uso de teclados especializados [14], hasta el uso de software especializado, que permiten incluso el acceso a la data en tiempo real [18]. Actualmente, el registro de un partido de fútbol típicamente involucra alrededor de ocho millones de datos. Debido al gran volumen de datos que genera la actividad, la información debe ser accedida y manipulada utilizando herramientas de *Big Data*. Esto impone una barrera tecnológica mínima necesaria para la utilización de la información generada para la toma de decisiones y análisis de datos en los clubes de fútbol.

Este gran volumen de datos entrega nuevas posibilidades en el mundo de la predicción de resultados deportivos. Un modelo base para esta materia es el que busca representar la conversión de goles de los equipos a través de procesos de Poisson, calibrando el modelo con variables como goles del equipo local, goles del equipo visitante y el efecto de la localía [3]. Sin embargo, este modelo y sus variantes utilizan datos agregados de los equipos. Esto tiene la limitante de que no se puede incorporar información individual de los jugadores para la predicción de resultados deportivos y, por otra parte, no permite individualizar el rendimiento de los jugadores para realizar *scouting*.

En este contexto, [2] estudia el proceso de detección de talentos que siguen 125 buscadores de talento (*scouts*) de distintas categorías en Países Bajos para poder determinar las características más importantes que, según ellos, son predictores del talento: a través de encuestas de respuestas abiertas, lo autores identifican los aspectos técnicos de los jugadores (37%) la categoría de atributos más importante, seguida de los aspectos tácticos y percepción-cognitiva (22%). Una vez definidas estas variables, se procede al análisis para determinar qué jugadores del mercado podrían ser atractivos para maximizar el rendimiento deportivo. En este ámbito, [6] propone medir el impacto que un jugador genera en un club en particular a través del modelamiento de un partido a través de un proceso estocástico Markoviano para medir dicho impacto. Nuestro trabajo puede ser visto como una continuación de la propuesta en [6], pero incluyendo inferencia Bayesiana para la calibración de los modelos propuestos, entre otros aspectos. Utilizando herramientas analíticas y datos de posesión de balón, [7] analiza líneas de pases que se generan cuando un jugador tiene el balón, clasificándolas como i) pases penetrantes, ii) pases de apoyo y iii) pases de seguridad, evaluando, en función del tiempo disponible las línea de pase disponibles. Otra área de aplicación, que no exploramos en esta revisión, es posible utilizar técnicas de analítica avanzadas para intentar predecir

lesiones en el fútbol. Las lesiones tienen un gran impacto en el deporte, por los costos de rehabilitación y la pérdida de ese elemento para la competencia durante un tiempo determinado. A modo de ejemplo, [15] propone clasificar a los jugadores en base a sus cargas de entrenamiento, utilizando redes neuronales convolucionales que utilizan datos en forma de series de tiempo multivariadas, con el objetivo de detectar potenciales ventanas de tiempo donde un jugador podría lesionarse y prevenir esta afección.

3. Modelo Matemático

En esta sección presentamos un modelo estocástico de desarrollo de un partido de fútbol. Como mencionamos en la Sección ??, el nivel de detalle del modelo debe permitir medir el efecto de un jugador en la alineación del equipo, de forma de poder estimar su impacto en la probabilidad de salir campeón de un torneo. Esto, de forma de facilitar, por ejemplo, la valoración de un precio justo para alguna transacción de jugadores entre dos clubes, o intentar predecir como sería el rendimiento de un futbolista que está dentro de los planes de contratación de un club. A modo de resumen, el modelo que proponemos visualiza el desarrollo de un partido como una cadena de Markov en tiempo discreto, donde un estado incluye, por ejemplo, donde se ubica el balón y que jugador lo tiene, además de incluir variables que detallan el estado global del partido. Las transiciones entre un estado y otro de la cadena están gatilladas por las acciones de los jugadores. En este trabajo consideraremos tan solo tres acciones: pases, regates y tiros. En particular, dicha transición ocurre producto de una combinación de eventos independientes: en primer lugar, un evento asociado al proceso de decisión del jugador (ámbito cognitivo-perceptual), quien selecciona que acción realizar (pase, regate o tiro), en función de variables contextuales del partido; y luego un evento relacionado con la ejecución de tal acción, es decir, si la realiza de éxito o no. Planteamos modelos de decisión logísticos para las probabilidades asociadas a estos eventos, las que dependen de parámetros asociados, entre otras cosas, a los jugadores de forma individual. Para estimar estos modelos adoptamos un enfoque de inferencia Bayesiano, el cual detallamos más adelante en esta sección, y que ilustramos en la Sección 3.3 utilizando los datos de la Premier League de Inglaterra, temporada 2017-2018.

3.1. Modelo Markoviano de un Partido de Fútbol

3.1.1. Información Preliminar.

La construcción del modelo depende fuertemente de los datos disponibles para su calibración. Considerando esto, comenzamos esta sección con una descripción de los datos a nivel de eventos.

Descripción de los datos. A grandes rasgos, los datos contienen información general asociada a las plantillas de los equipos, eventos de cada uno de los partidos, e información del rendimiento deportivo de los jugadores. Para cada uno de los partidos del torneo se tiene el detalle de cada evento deportivo que ocurre durante el encuentro, esto es: pases, tarjetas, tiros, faltas, regates, detenciones, entre otros. En promedio ocurren casi 1.500 eventos relacionados al balón por partido, considerando las cinco ligas más grandes de Europa. Un evento se compone de múltiples características: la zona del campo de juego donde ocurre (punto (x_i, y_i) del evento i geo-referenciado en la cancha), el tiempo relativo al inicio del partido cuando ocurre el evento, el jugador asociado a este, si realiza la acción de forma exitosa o no, y si existe un segundo jugador asociado. Con este nivel de datos desagregados puede reconstruir el desarrollo del partido a nivel de la estructura y progreso del juego y por ende, el del campeonato completo.

Resumen del modelo, estados y transiciones Representamos el desarrollo del partido como una cadena de Markov en tiempo discreto $\{X_n : n \in \mathbb{N}\}$, donde X_n representa un estado de desarrollo del partido tras n transiciones o períodos. Si bien no modelamos directamente la duración de un período, si controlamos el número de períodos que contiene un partido mediante una discretización de la duración del mismo (detalles más abajo). Para que un proceso estocástico como $\{X_n : n \in \mathbb{N}\}$ sea una cadena de Markov es necesario que X_n contenga toda la información necesaria para caracterizar (probabilísticamente) la evolución futura del proceso. En nuestro modelamiento, las transiciones entre estados de la cadena de Markov son gatilladas por las decisiones tomadas por el jugador que tiene el balón al comienzo de un periodo, y por el subsecuente éxito o fracaso de dicha acción. Modelamos la decisión del jugador, y el resultado de la ejecución como eventos independientes, condicional en el estado del sistema.

3.1.2. Estados de la cadena de Markov.

En este trabajo, utilizaremos la siguiente representación, la cual implícitamente asumiremos otorga la condición de Markov al desarrollo del partido:

$$X_n = (X_n^1, X_n^2, X_n^3, X_n^4), \quad n \in \mathbb{N},$$

donde

- X_n^1 representa el jugador que tiene el balón al comienzo del período: las transiciones entre estados, se gatillan por las acciones tomadas por este jugador, y el éxito en su ejecución. Ver más detalles, abajo.
- X_n^2 representa la zona de la cancha donde se encuentra el jugador: para evitar trabajar con una representación continua (en dos dimensiones) de la ubicación del balón, lo que implica un grado de sofisticación mayor para determinar las transiciones espaciales de la cadena, dividimos el campo de juego en 12 zonas, producto de dividir la cancha horizontal/verticalmente en 4/3 zonas. Para más detalles respecto a esta discretización, ver [12].
- X_n^3 representa el tiempo transcurrido desde el inicio del partido: tal como en el caso anterior, evitamos trabajar en tiempo continuo utilizando una discretización del tiempo de juego. En particular, supusimos que cada acción toma un tiempo determinista Δ , por lo que el tiempo de juego periodo a periodo avanza en esa cantidad. El número total de períodos en un partido se ajustó considerando la historia de encuentros entre los equipos. Para más detalles, ver [12]¹.
- X_n^4 contiene información adicional asociada al contexto general del partido: en particular, incluimos en esta información el resultado parcial del partido, es decir, si el equipo que tiene el balón está ganando, empatando o perdiendo.

Información adicional que no cambia durante el partido, como por ejemplo el estado de localía o visita, también es considerada, pero no como parte del estado.

3.1.3. Transiciones de estados

La transición de un estado a otro involucra, en orden cronológico, los siguientes pasos.

1. Primero, el jugador decide que acción ejecutar.
2. Segundo, dependiendo de la acción, un jugador secundario es seleccionado para participar de la acción: que en el caso de pases, se selecciona un jugador del mismo equipo que recibirá el pase, junto a una zona de

¹Notamos que este enfoque es equivalente remover X_n^3 de la descripción de estado, e incluir una dependencia en el periodo n en las probabilidades de transición. Preferimos el enfoque actual, que facilita la descripción del modelo.

recepción del pase - la selección se realiza considerando la frecuencia con la que pases con esa zona de origen se realizan a esos jugadores/zonas; para el caso de regates, se selecciona un jugador de equipo rival, considerando la frecuencia con la cual dichos jugadores se ven involucrados en regates en esa zona del campo; en el caso de tiros, el jugador secundario es siempre el arquero del equipo rival².

3. Tercero, se determina si la acción es exitosa o no. Finalmente, si la acción no es exitosa, se determina qué jugador del equipo rival recupera el balón; exceptuando los pases, este jugador es el jugador secundario asociado a la jugada; en el caso de un pase, el jugador se elige (de manera frecuentista) utilizando la frecuencia con la cual los jugadores interceptan pases, dependiendo de la zona de la cancha involucrada.

De la descripción anterior, vemos que hay dos eventos cuyas probabilidades asociadas falta especificar: la selección de la acción a realizar, y el éxito de dicha acción. A continuación, modelamos estas probabilidades adoptando un enfoque paramétrico, que luego incorporaremos dentro de un enfoque de inferencia bayesiano.

- **La decisión del jugador.** En términos de las acciones a realizar, para la construcción de este modelo se consideran los pases y regates, que representan más del 75% de los eventos que ocurren en un partido de fútbol y los tiros que dan pie a representar los goles, el evento con mayor importancia de este deporte [12]. El resto de los eventos corresponden a interrupciones y reposiciones de balón, los cuales no son abordadas en este modelo. Para modelar la decisión del jugador utilizaremos modelos de decisión discreta (ver, e.g., [22]), capaces de reflejar el hecho que los jugadores no siempre toman la misma decisión cuando se enfrentan a situaciones similares, y si bien tienden a privilegiar ciertas acciones sobre otras, en general existe un grado de aleatoriedad en la decisión, necesaria para evitar que los rivales puedan predecir sus acciones. En particular, en este trabajo asumiremos que los jugadores utilizan un modelo de decisión Logit multinomial [22].

²Notamos que las probabilidades de selección de jugadores secundarios se calibran usando una lógica frecuentista, por lo que no se incluyen dentro del marco de inferencia bayesiana a detallar mas adelante en esta sección. Este supuesto se realiza para reducir la complejidad de la tarea de inferencia.

Dado un estado X_n al comienzo del periodo n , consideramos la variable aleatoria $Y_n \in \{p \text{ (pase)}, t \text{ (tiro)}, r \text{ (regate)}\}$. Supondremos que

$$\begin{aligned} \mathbb{P}\{Y_n = p|X_n\} &= \frac{\exp(Z_p(X_n))}{1 + \exp(Z_p(X_n)) + \exp(Z_t(X_n))}, \\ \mathbb{P}\{Y_n = t|X_n\} &= \frac{\exp(Z_t(X_n))}{1 + \exp(Z_p(X_n)) + \exp(Z_t(X_n))}, \\ \mathbb{P}\{Y_n = r|X_n\} &= \frac{1}{1 + \exp(Z_p(X_n)) + \exp(Z_t(X_n))}, \end{aligned}$$

donde

$$Z_i(x_n) := \beta_{i,0} + \beta_{i,X_n^1} + \beta_{i,X_n^2} + \beta_{i,X_n^3} + \beta_{i,X_n^4}, \quad i \in \{p, t\} \quad (1)$$

Esto es, la decisión de que acción realizar depende de una *propensión* de cada jugador a realizar cada acción, la que es afectada transversalmente por la zona de la cancha donde se encuentra el balón, del tiempo transcurrido del partido, y el resultado parcial. Notamos que el factor de localía se incluye en el termino inicial $\beta_{i,0}$, que considera dos valores posibles, dependiendo el jugador se encuentra en un equipo que juega en calidad de local o visita (no se registran partidos en terreno neutral en nuestros datos).

- El resultado de la acción.** Modelamos el resultado de una acción mediante un modelo logístico, donde juegan un rol distintos factores dependiendo de las características de la acción. En general, el modelo utilizado es:

$$\mathbb{P}(\text{Exito}|A) = \frac{\exp(W(A))}{1 + \exp(W(A))},$$

donde A contiene las características de la acción (jugador, tiempo, zona, etc.). A continuación describimos el modelo asociado a cada una de las acciones consideradas.

Pases. En el caso de pases tenemos que $A = (X_n, J, Z)$ donde J representa el jugador al cual el pase esta dirigido, y Z la zona del campo donde se recibirá el pase. En este caso tenemos que

$$W(A) := \alpha_{p,0} + \alpha_{p,X_n^1}^o + \alpha_{p,J}^d + \alpha_{p,X_n^2,Z} + \alpha_{p,X_n^3} + \alpha_{p,X_n^4}. \quad (2)$$

Vemos que el éxito del pase depende tanto del ejecutante como del receptor del pase, los sectores del campo involucrados, y las condiciones del encuentro.

Tiros. En el caso de tiros tenemos que $A = (X_n, J)$ donde J representa el arquero del cuadro rival. En este caso tenemos que

$$W(A) := \alpha_{t,0} + \alpha_{t,X_n^1}^o + \alpha_{t,J}^d + \alpha_{t,X_n^3} + \alpha_{t,X_n^4}. \quad (3)$$

Entonces, vemos que la probabilidad de conversión del tiro depende tanto del ejecutante como del arquero rival, y las condiciones del encuentro.

Regates. En el caso de regates tenemos que $A = (X_n, J)$ donde J representa el jugador al cual se encara durante el regate. Tal como en el caso anterior, tenemos que

$$W(A) := \alpha_{r,0} + \alpha_{r,X_n^1}^o + \alpha_{r,J}^d + \alpha_{r,X_n^3} + \alpha_{r,X_n^4}. \quad (4)$$

Entonces, vemos que la probabilidad de sobrepasar al jugador rival depende tanto del ejecutante como del rival, y las condiciones del encuentro.

3.2. Inferencia Bayesiana del Modelo

De la sección anterior, vemos que la selección probabilista de la acción a realizar queda determinado por el vector de parámetros $\beta := (\beta_t, \beta_p)$ aludidos en la ecuación (1), donde $\beta_i := (\beta_{i,j} : j \in J)$ donde el conjunto J incluye todos los jugadores, zonas del campo, periodos, resultados parciales, y situaciones de localía posibles, $i \in \{t, p\}$. Notamos que, así como no es necesario especificar parámetros asociados a la acción de regate (esto, pues la magnitud de parámetros asociados a tiro y pase se entienden como expresados relativos a aquellos - no especificados - asociados a regate), para la correcta identificación del modelo tampoco es necesario especificar parámetros para todas las opciones en J , y es posible expresar estos parámetros relativos a un jugador, un tiempo, una zona, y un resultado en particular. De la misma forma, la probabilidad de éxito de la acción seleccionada queda determinada por el vector de parámetros $\alpha := (\alpha_t, \alpha_p, \alpha_r)$ aludidos en las ecuaciones (2)(3) y (4), donde $\alpha_i = (\alpha_{i,j} : j \in J_i)$ donde el conjunto J_i incluye todas las situaciones en las que se puede realizar la acción $i \in \{t, p, r\}$. Al igual que en el caso anterior, no todos los coeficientes necesitan ser especificados para la correcta identificación del modelo.

Incluso con las consideraciones mencionadas arriba, el número de parámetros necesario identificar para especificar el modelo es grande. Más importante aun, muchas de las acciones que son directamente observables en los datos dependen de la interacción entre múltiples parámetros, por lo que la inferencia de tipo frecuentista es compleja. Con esto en mente, proponemos un esquema de inferencia bayesiana. Esto es, supondremos que los parámetros (α, β)

son a priori variables aleatorias: cada realización de esos parámetros aleatorios determina un modelo distinto. Sin embargo, supondremos que la realización toma lugar antes del comienzo del torneo (i.e. antes de que se recojan los datos), y se mantiene constante a través del torneo, por lo que los datos son originados condicionales en una realización en particular de los parámetros; el trabajo de inferencia consiste en estimar cual es esta realización.

Dado el modelo probabilista y de generación de datos, la estimación de los parámetros consiste en el cálculo de la distribución a posterior de estos, condicional en los datos observados. Esto es, si (α, β) distribuyen *a priori* de acuerdo a una densidad $f(\cdot)$, y $L(\text{data}|\alpha, \beta)$ representa la verosimilitud asociada a los datos observados, condicional en los parámetros (α, β) , entonces la densidad *posterior* de los parámetros condicional en los datos, $f(\alpha, \beta|\text{data})$, es tal que

$$f(\alpha, \beta|\text{data}) \propto L(\text{data}|\alpha, \beta) \cdot f(\alpha, \beta) \quad (5)$$

Este esquema de inferencia mostrado en (5) se apoya fundamentalmente en el teorema de Bayes, el que en el caso que el modelo es identificable a partir de los datos, garantiza la consistencia de los estimadores. En el caso de nuestro modelo, esta distribución a posterior no es calculable de forma analítica, y debe ser estimada numéricamente. Para esto utilizamos el enfoque de Monte Carlo Markov Chain [13], el que aproxima la distribución posterior utilizando técnicas de simulación. El método queda definido, entre otras cosas, por la distribución a priori dada a los parámetros. Para esto, en nuestros experimentos numéricos, presentados en la Sección 3.3, utilizamos un prior normal, independiente para cada parámetro. Esto es, suponemos que

$$f(\alpha, \beta) = \phi(\alpha, \beta, \mu, \Sigma),$$

donde μ y σ son un vector y una matriz diagonal, positiva, de las dimensiones apropiadas, ϕ representa la densidad de un vector normal multivariado, también de las dimensiones correctas. Para mas detalles respecto a la selección de las distribuciones a priori de los parámetros (esto es, los valores de μ y Σ utilizados), ver [12].

3.3. Limitaciones del modelo

El modelo propuesto cuenta con variadas limitaciones, algunas originadas por los datos utilizados, y otras por los supuestos simplificadores que nos permiten representar un partido de fútbol como una cadena de Markov. Desde el punto de vista de los datos, es importante considerar que estos solo se refieren a la información del jugador o los jugadores que participan directamente con

el balón, pero no se cuenta con información del resto de los jugadores que no participan activamente con el balón (que se puede argumentar es tan importante como el resto), por lo que no se captura toda la información de que está ocurriendo en el partido. Por otra parte, este modelo es una representación simplificada de un partido de fútbol y hay ciertos elementos que no son considerados, como por ejemplo, la formación de cada equipo (lo que puede cambiar la propensión de un jugador a tomar una u otra acción), o de forma similar, la posición de cada jugador dentro del terreno de juego (considere el comportamiento de un mismo jugador cuando juega como central versus volante). Estas limitaciones serán discutidas más en detalle en la Sección 4.2

4. Resultados

En esta sección ilustramos la aplicación del enfoque propuesto, utilizando los datos de la temporada 2017-2018 de la Premier League de Inglaterra, de la cual se disponen datos a nivel eventing [16]. Utilizando la dinámica del modelo propuesto, programamos rutinas de simulación capaces de muestrear el desarrollo de un partido de acuerdo a las reglas probabilistas de nuestro modelo, posibilitando la simulación de torneos completos y la realización de ejercicios de sensibilidad como por ejemplo, estudiar el efecto de intercambiar un par de jugadores entre equipos rivales. Todas las rutinas fueron programadas en el lenguaje Python [23]; las rutinas de inferencia bayesiana fueron programadas desde Python usando el lenguaje de modelamiento Stan [21], a través de la librería PyStan.

4.1. Comparación con otro modelo

Comenzamos el análisis realizando una simulación del primer partido del torneo, entre los equipos Arsenal y Leicester City, cuyo resultado fue 4-3 a favor del local. La Tabla 1 muestra estadísticas agregadas de 10.000 simulaciones de dicho encuentro.

Equipo	Promedio eventos	Promedio goles	Desv. estándar goles
Arsenal	794,78	3,14	1,23
Leicester City	657,02	2,66	1,10

Tabla 1: Promedio de eventos y goles: Arsenal vs Leicester City.

Nuestra simulación concluye que el Arsenal anota (en promedio) cerca de

1/2 gol más que el Leicester cuando estos dos equipos se encuentran. En ese sentido, si bien el resultado real del encuentro fue 4-3 a favor del local, nuestro modelo indica que la victoria del Arsenal es una realización de múltiples posibilidades, y que no existe un dominio total sobre el rival, como lo ilustra la desviación estándar del número de goles anotados.

Si bien nuestro modelo está diseñado para representar de buena forma los sucesos ocurridos durante un partido, y no pronosticar el resultado final, el que puede muchas veces estar desconectado del desarrollo de un encuentro (el fútbol es uno de los deportes masivos mas difícil de predecir debido a esta desconexión, y al relativo bajo número de anotaciones por encuentro [1]) es interesante realizar la comparación contra modelos alternativos que si tienen ese propósito (pero que no sirven para apoyar labores de scouting, por lo mismo). Con esto en mente, comparamos nuestro modelo contra una variación del modelo de Poisson propuesto por [3]. Dicho modelo se alimenta solamente de los resultados finales de los partidos del torneo, más las condiciones de localía. Con esto, el modelo calibra **tasas de ataque y defensa** para cada equipo, más un factor de localía; para simular un partido, estas tasas se utilizan para calcular tasas de procesos de Poisson (uno para cada equipo), los que representan los procesos de anotación de goles de cada equipo; la simulación del resultado del partido se genera a través del muestro de dos variables aleatorias distribuidas Poisson.

Para comparar modelos consideramos la métricas de **acierto de resultado** (local, empate, o visita) y **acierto diferencia de goles** (goles del local menos goles de visitante) que indican el porcentaje del partidos (calculado sobre base a 10.000 simulaciones del torneo) en los que un modelo genera un resultado/diferencia de goles que coincide con el resultado/diferencia real del partido. La Tabla 2 muestra los resultados de la comparación de los modelos.

Parámetros	C. Markov		Poisson	
	1-19	20-38	1-19	20-38
% acierto al resultado	39,09	39,12	44,28	44,28
% acierto diferencia de goles	18,38	19,26	17,64	18,54

Tabla 2: Rendimiento de los modelos de cadena de Markov y Poisson para 10.000 iteraciones según diferentes parámetros.

En la Tabla 2 los modelos son primero calibrados utilizando los partidos de la primera rueda (fechas 1-19), y simulando las partidos de la segunda rueda (fechas 20-38) para el calculo de las métricas, y luego invirtiendo los roles de la primera y segunda ruedas. En ambos casos notamos que nuestro modelo tiene un rendimiento aproximado del 39% en contraste con el 44% del

modelo de Poisson. Es decir, este último es capaz de predecir en un 44 % el resultado (local, empate o visita) de los partidos, calibrando el modelo sólo con la mitad de los datos (fecha 20 a 38). Por otro lado, vemos que nuestro modelo predice la diferencia de goles en un 18-19 % de los casos, superando levemente el resultado del modelo de Poisson. Este resultado es alentador, considerando que el modelo de Poisson esta calibrado explícitamente para replicar diferencias de goles. En ese sentido, nuestro modelo se comporta en forma comparable al modelo estándar para predecir resultados agregados, pero tiene la capacidad adicional de permitir hacer análisis a nivel de jugador, que es lo procedemos a ilustrar a continuación.

4.2. Análisis de Sensibilidad: Intercambio de Jugadores

Para ilustrar las posibilidades de análisis que permite el modelo, en esta sección consideramos el ejercicio donde intercambiamos dos jugadores entre equipos rivales, y estudiamos el efecto de dicho intercambio en la probabilidad de dichos equipos de ganar el campeonato³. Los pares de jugadores fueron escogidos por jugar en una misma posición y en base a un análisis cualitativo realizado por los autores. En total consideramos 6 dichos intercambios. Para comenzar, simulamos el torneo 100, 1.000, y 10.000 veces (sin intercambio alguno) para aproximar la probabilidad base de cada equipo de salir campeón. Estas estimaciones se muestran en la Tabla 3, junto al resultado real (en puntos) del torneo.

Una observación destacable desde la tabla 3 son las probabilidades de salir campeón de los equipos de Manchester. Nuestro modelo indica que el M. United tiene un “mejor plantel” relativo al M. City, por lo que la probabilidad de salir campeón es 10 % mayor. Por otro lado, los datos indican que durante esa temporada el M. United hizo cerca de 6 goles menos de los que debiese haber hecho (considerando *expected goals*), y por otro lado el M. City hizo 17 goles más de lo que debiese haber hecho. En forma más coloquial, los jugadores del M. United tuvieron un rendimiento bajo el promedio esperado y los delanteros del M. City tuvieron un rendimiento por sobre lo esperado, en términos de conversión de goles basado a la calidad de las oportunidades generadas (*expected goals*). Esto es parte de lo estocástico que puede ser el fútbol.

³Vale la pena notar que nuestro modelo ignora ciertas consideraciones estratégicas, y supone que los equipos siempre ponen el mismo esfuerzo en cada encuentro. En ese sentido, el modelo no considera el objetivo de los equipos, y el análisis presente no incorpora de forma especial el objetivo de ganar el campeonato, por lo que otras medidas, como por ejemplo la probabilidad de descender de categoría, también pueden ser analizadas.

Tabla de posiciones real		Probabilidad de campeón		
Equipo	Ptos.	100 its.	1.000 its.	10.000 its.
		%	%	%
Manchester City	100	25,00	20,34	20,50
Manchester United	81	27,00	31,11	31,20
Tottenham Hotspur	77	11,00	18,58	18,52
Liverpool	75	11,00	9,32	9,29
Chelsea	70	8,00	4,81	5,35
Arsenal	63	2,00	1,23	1,24
Burnley	54	6,00	4,16	4,12
Everton	49	0,00	0,93	0,81
Leicester City	47	10,00	7,92	7,45
Newcastle United	44	0,00	0,21	0,19
Crystal Palace	44	0,00	0,00	0,01
Bournemouth	44	0,00	0,17	0,11
West Ham United	42	0,00	0,97	0,67
Watford	41	0,00	0,00	0,00
Brighton & H. A.	40	0,00	0,00	0,04
Huddersfield Town	37	0,00	0,00	0,00
Southampton	36	0,00	0,00	0,02
Swensea City	33	0,00	0,00	0,00
Stoke City	33	0,00	0,25	0,26
West Bromwich A.	31	0,00	0,00	0,04

Tabla 3: Probabilidad de salir campeón de los equipos, luego de 100, 1.000 y 10.000 iteraciones del torneo.

1. **Harry Maguire por Rob Holding (defensas).** El primer intercambio de jugadores que se realiza es el de Harry Maguire, jugador perteneciente al Leicester City por Rob Holding, perteneciente al equipo Arsenal. Ambos son jugadores que juegan en la posición de defensas centrales y participaron en 3.420 minutos y 820 minutos respectivamente. La Figura 1 muestra el desempeño de ambos considerando los parámetros: porcentaje de duelos defensivos ganados, porcentaje de duelos aéreos ganados, porcentaje de pases cortos correctos, porcentaje de pases largos correctos y porcentaje de minutos jugados en relación a todo el torneo. Si bien, ambos son jugadores relativamente parejos en relación al desempeño que tuvieron durante el campeonato, el jugador Harry Maguire

le saca ventajas a Holding en el porcentaje de duelos aéreos ganados, porcentaje de duelos defensivos ganados y estuvo en cancha más de cuatro veces los minutos en los que participó Rob Holding. Por otro lado, Holding, es levemente superior en los parámetros de pases cortos y pases largos, con un acierto cercano al 80 % y 60 % respectivamente.

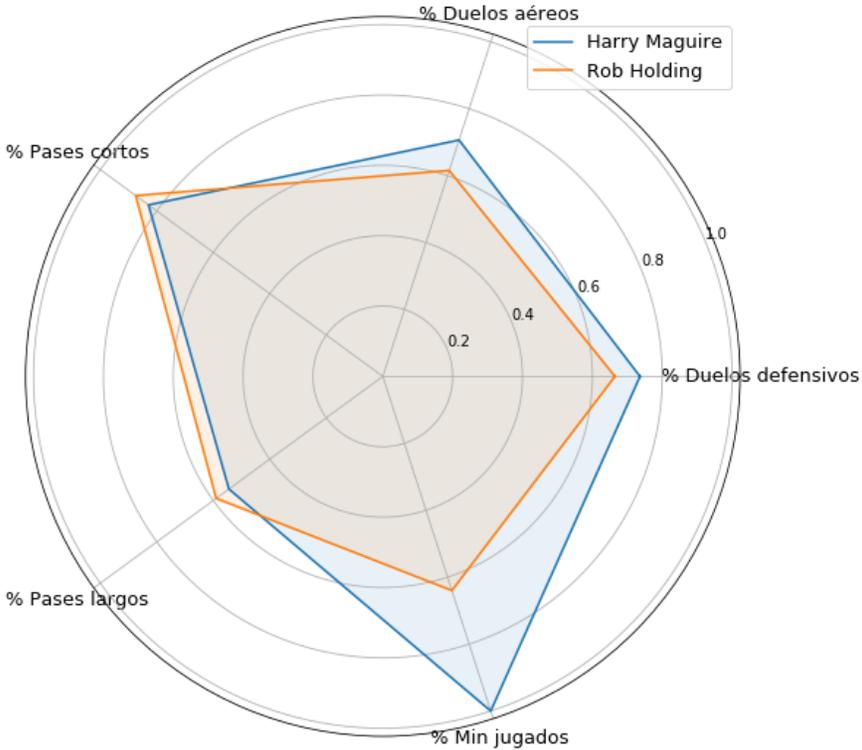


Figura 1: Comparación de rendimiento defensivo del jugador Harry Maguire (Leicester City) y Rob Holding (Arsenal).

La Tabla 4 muestra la variación en las probabilidades de ganar el campeonato para el Arsenal y el Leicester producto de este intercambio; para ver el detalle completo de las nuevas probabilidades de campeonar de todos los equipos, ver [12].

Se observa que con la salida de Harry Maguire y la inclusión de Holding en el Leicester City, las probabilidades de campeonar de este equipo bajan en 0,75 puntos porcentuales, de forma inversa la inclusión de Maguire en el Arsenal hace que este equipo se refuerce de mejor manera y sus probabilidades aumentan en 1,16 puntos porcentuales.

Tabla 4: Probabilidades de campeonar intercambiando a Maguire por Holding.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Arsenal	1,24	2,40	+ 1,16
Leicester City	7,45	6,70	- 0,75

2. **David Silva por Dele Alli.** Los resultados del intercambio de David Silva (2.438 min. jugados) jugador del equipo campeón Manchester City por el jugador del Tottenham Hotspur Dele Alli (2.971 min. jugados), se muestran a continuación. En primer lugar, se analiza el desempeño de ambos jugadores en el torneo, tal y como se muestra en la Figura 2.

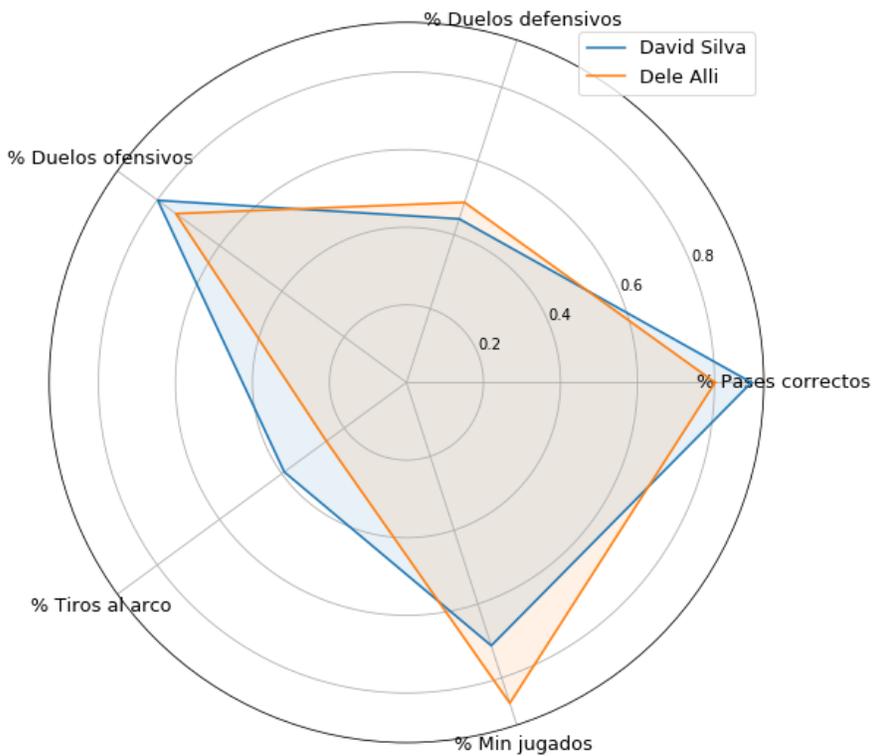


Figura 2: Comparación de rendimiento en el mediocampo del jugador David Silva (Manchester City) y Dele Alli (Tottenham Hotspur).

Con respecto a la variación en la probabilidad de campeonar, como se muestra en la Tabla 5 el Manchester C. disminuye en 9,90 puntos porcentuales su probabilidad de campeonar, mientras que el Tottenham con la inclusión de D. Silva aumenta sus probabilidades en 20,58 puntos por-

centuales. Este es un buen ejemplo, para determinar que Silva podría ser un gran refuerzo en un equipo como el Tottenham dado que con su inclusión las probabilidades de campeónar de este último equipo aumentan considerablemente: para ver el detalle completo de las nuevas probabilidades de campeónar de todos los equipos, ver [12].

Tabla 5: Probabilidades de campeónar intercambiando a Silva por Alli.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Manchester City	20,50	10,60	- 9,90
Tottenham Hotspur	18,52	39,10	+ 20,58

3. **Kevin de Bruyne por Mark Noble (mediocampistas).** Los resultados del intercambio entre Kevin De Bruyne (3.084 min. jugados) jugador del equipo Manchester City y Mark Noble (2.404 min. jugados), histórico medicampista del West Ham United desde el año 2004, se detallan a continuación. En la Figura 3 se realiza una comparación del desempeño durante el torneo de estos dos jugadores.

Con respecto a la variación en la probabilidad de campeónar, como se muestra en la Tabla 6 el Manchester C. disminuye en 5,50 puntos porcentuales su probabilidad de campeónar, mientras que el West Ham U. con la incorporación de K. De Bruyne aumenta sus probabilidades en 1,83 puntos porcentuales. Aquí se observa lo importante del proceso de *scouting*, ya que la inclusión de un jugador puede incluso cuadruplicar las probabilidades de campeónar, tal como lo sería K. De Bruyne en el West Ham United: para ver el detalle completo de las nuevas probabilidades de campeónar de todos los equipos, ver [12].

Tabla 6: Probabilidades de campeónar intercambiando a De Bruyne por Noble.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Manchester City	20,50	13,70	- 5,50
West Ham United	0,67	2,50	+ 1,83

4. **Romelu Lukaku por Álvaro Morata (delanteros).** Los resultados del intercambio entre el traspaso más caro del torneo, Romelu Lukaku (2.869 min. jugados), jugador del equipo Manchester United y Álvaro Morata (2.068 min. jugados), jugador del Chelsea y de la selección española, se detallan a continuación. En la Figura 4 se muestra una comparación del desempeño durante el torneo entre estos dos jugadores. Lukaku es superior en todos los parámetros a analizar con excepción de

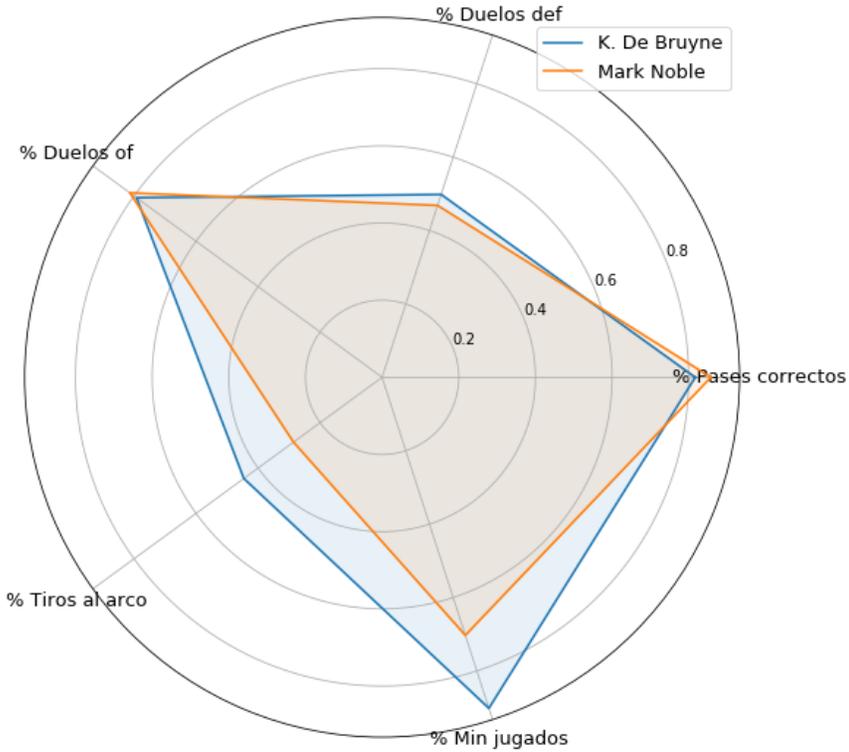


Figura 3: Comparación de rendimiento en el mediocampo del jugador Kevin De Bruyne (Manchester City) y Mark Noble (West Ham United).

los pases correctos, donde Morata tiene una mayor efectividad.

Tal como lo muestra la Figura 4 la salida de Lukaku y la integración de Morata en el Manchester United beneficia al Chelsea, ya que Lukaku es un delantero con mejor rendimiento. En la Tabla 7 se ve que el intercambio entres estos dos ocasiona que el Manchester United se vea perjudicado, disminuyendo en 13,10 puntos porcentuales su probabilidad de campeonar, en contraposición al Chelsea que se vería beneficiado aumentándola en 5,87 puntos porcentuales: para ver el detalle completo de las nuevas probabilidades de campeonar de todos los equipos, ver [12].

Tabla 7: Probabilidades de campeonar intercambiando a Lukaku por Morata.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Manchester United	31,20	18,10	- 13,10
Chelsea	5,53	11,40	+ 5,87

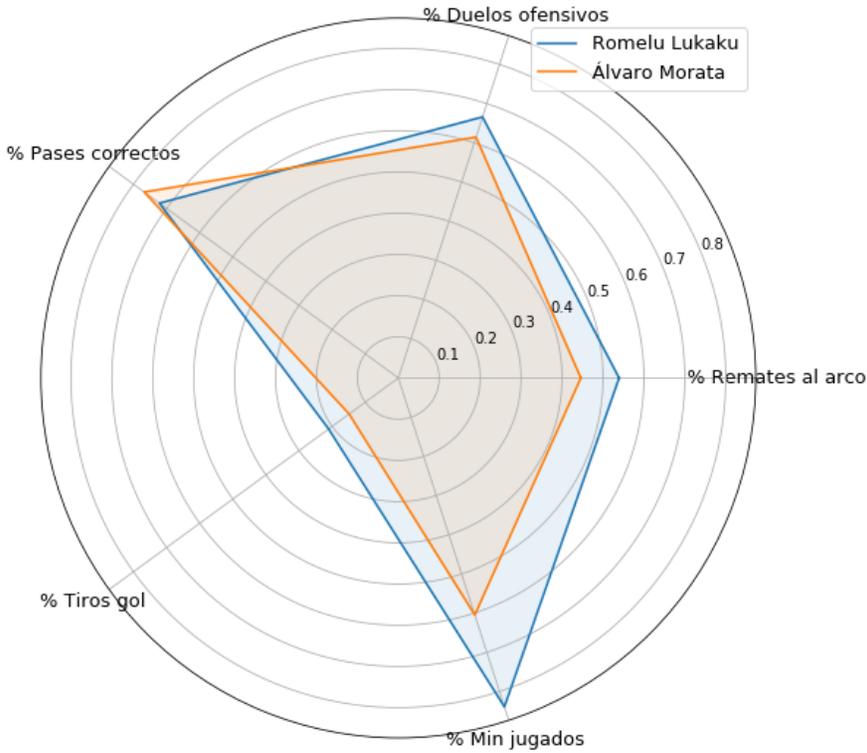


Figura 4: Comparación de rendimiento ofensivo del jugador Romelu Lukaku (Manchester United) y Álvaro Morata (Chelsea).

5. **Sergio Agüero por Alexandre Lacazette (delanteros).** Los resultados del intercambio de Sergio Agüero (1.968 min. jugados), goleador histórico del equipo campeón Manchester City (21 goles en el torneo) y el jugador del Arsenal Alexandre Lacazette (2.211 min. jugados), se muestran a continuación. En primer lugar, se revisa el desempeño de ambos jugadores en el torneo, en el gráfico de la Figura 5.

Con respecto a la variación en la probabilidad de campeonar, como se muestra en la Tabla 8 el Manchester City disminuye en 6,80 puntos porcentuales su probabilidad de campeonar, mientras que el Arsenal con la incorporación de Sergio Agüero aumenta sus probabilidades en 2,46 puntos porcentuales. Si bien Lacazette tiene un mayor porcentaje de tiros al arco, Agüero hizo siete goles más en siete partidos menos jugados, lo que lo convierte en un atacante muy efectivo con 0,84 goles por partido, justo detrás del goleador del torneo Mohamed Salah, mientras que Alexandre Lacazette tiene 0,44 goles por partido, casi la mitad: para ver

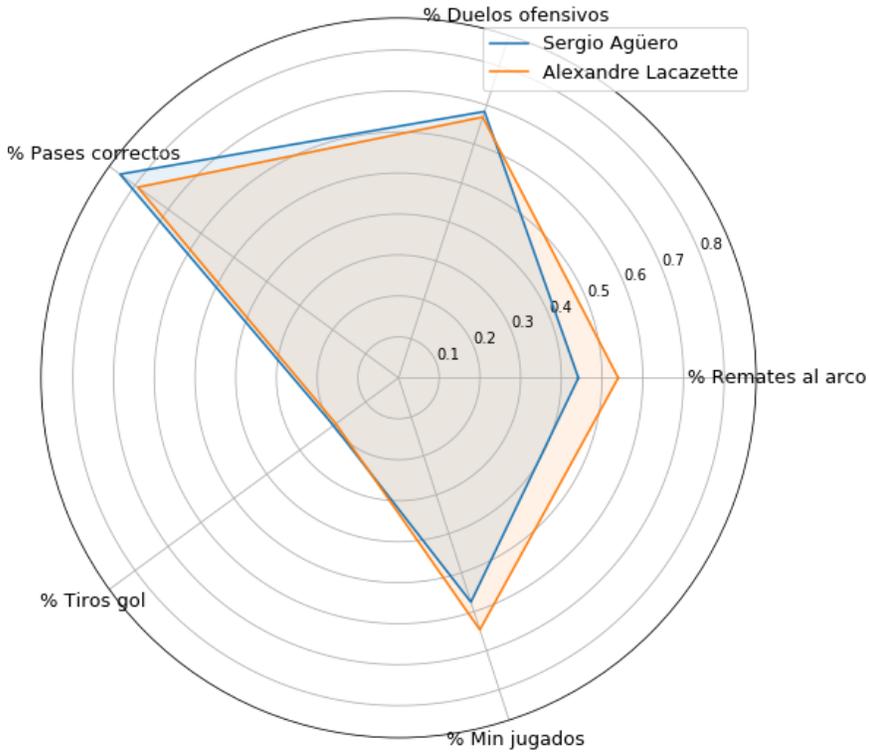


Figura 5: Comparación de rendimiento ofensivo del jugador Sergio Agüero (Manchester City) y Alexandre Lacazette (Arsenal).

el detalle completo de las nuevas probabilidades de campeonar de todos los equipos, ver [12].

Tabla 8: Probabilidades de campeonar intercambiando a Agüero por Lacazette.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Manchester City	20,50	13,70	- 6,80
Arsenal	1,24	3,70	+ 2,46

6. **Mohamed Salah por Mame Diouf (delanteros).** Los resultados del intercambio entre el goleador del torneo Mohamed Salah (2.922 min. jugados), jugador del Liverpool y Mame Diouf (2.601 min. jugados), histórico atacante del Stoke City desde el año 2014 al 2020, se detallan a continuación. En la Figura 6 se realiza una comparación del desempeño durante el torneo de estos dos jugadores. En donde se observa claramente que Salah es un jugador mucho más completo en relación a estos

parámetros, ya que supera en todo nivel a M. Diouf.

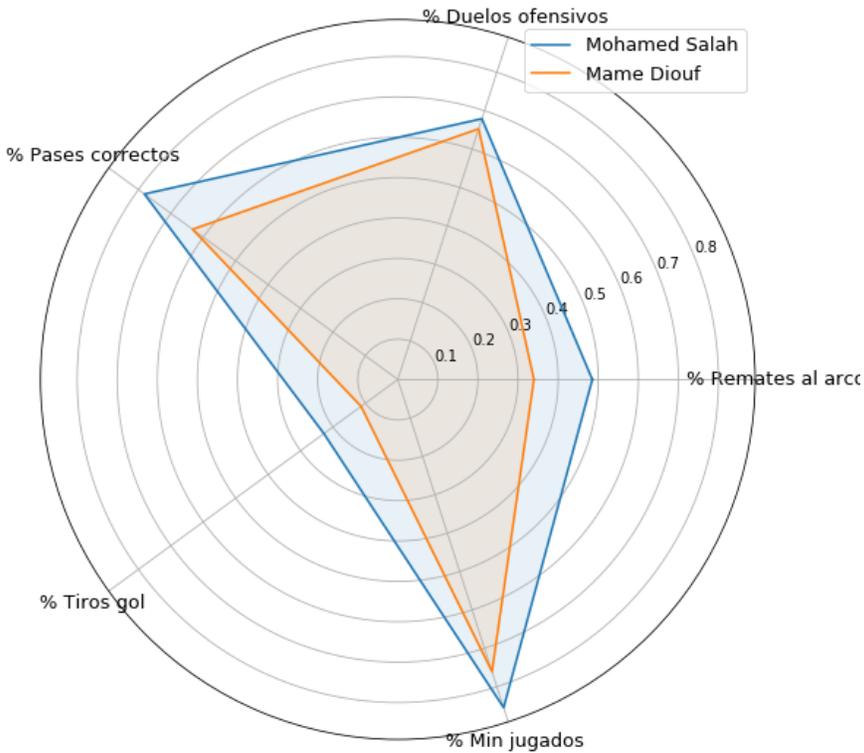


Figura 6: Comparación de rendimiento ofensivo del jugador Mohamed Salah (Liverpool) y Mame Diouf (Stoke City).

Con respecto a la variación en la probabilidad de campeonar, como se muestra en la Tabla 9 el Liverpool disminuye en 8,89 puntos porcentuales su probabilidad de campeonar, mientras que el Stoke City con la incorporación de Mohamed Salah aumenta sus probabilidades en 6,34 puntos porcentuales. La incorporación del goleador (32 goles) del campeonato en un equipo que terminó en el puesto 19° con tan solo 35 goles convertidos, casi los mismos goles convertidos por M. Salah en 36 partidos hace que las probabilidades de campeonar aumenten casi 25 veces: para ver el detalle completo de las nuevas probabilidades de campeonar de todos los equipos, ver [12].

Tabla 9: Probabilidades de campeón intercambiando a Salah por Diouf.

Equipo	Probs. sin cambio	Probs. con cambio	Delta
Liverpool	9,29	0,40	- 8,89
Stoke City	0,26	6,60	+ 6,34

5. Discusión y conclusiones

La recolección exhaustiva de datos a nivel eventing en la mayoría de las ligas profesionales del planeta plantea un desafío de Big Data en este deporte: la creación de herramientas analíticas para el apoyo a la toma de decisiones. Por otro lado, la literatura actual en fútbol analytics plantea modelos que no son adecuados para apoyar labores de scouting, puesto que se basan en el comportamiento agregado de los equipos, sin utilizar estadísticas individuales de los jugadores.

Con esto en mente, en este trabajo planteamos un enfoque de modelamiento que permite medir el rendimiento de los futbolistas a nivel individual y cuantificar el impacto que la inclusión de un futbolista genera en un determinado equipo, para luego ser utilizado en el proceso de *scouting*. Nuestro enfoque plantea la calibración del modelo utilizando inferencia bayesiana, y estima la capacidad cognitiva-perceptual (toma de decisiones) y capacidad técnica (ejecución de una acción) de cada jugador, permitiendo también ver el impacto de la inclusión de un futbolista en un equipo en particular medido como la probabilidad de ser campeón de dicho equipo. Este tipo de herramientas permite apoyar la toma de decisiones sobre cuáles son los futbolistas que más aportan al equipo, en base al estilo de juego de este y cómo este jugador impacta en a nivel agregado. Se observa que la *salida* de jugadores de buen rendimiento en equipos grandes impacta en forma significativa y negativa al equipo, mientras que su inclusión en equipos de menor categoría impacta en forma positiva pero no en la misma proporción que el equipo afectado. Esto beneficia más a los equipos que disputan el torneo en forma directa al equipo que se le sustrajo un buen elemento (ejemplo: reemplazo de (3) De Bruyne por Noble o (4) Lukaku por Morata). Según la experiencia en clubes de los autores, al momento de contratar un jugador se revisan 3 componentes: 1) reporte cuantitativo del jugador, 2) reporte cualitativo en base a videos y 3) situación contractual. Es posible ligar los resultados de este modelo al reporte cuantitativo que el tomador de decisiones requiere y complementar con los otros dos elementos. Además, esta herramienta sirve para acotar la búsqueda

de jugadores y utilizar menos horas pero con mayor foco en la revisión de videos de futbolistas.

El modelo propuesto presenta variadas limitaciones. Por un lado, el modelamiento solo describe el movimiento del balón sobre el campo de juego, ignorando el juego que se realiza sin el balón; esto se debe al tipo de dato con el que se trabaja (el que representa lo que ocurre con los jugadores que tienen el balón pero no con aquellos jugadores que no tienen el balón en un instante en particular); sin embargo, tecnologías para el seguimiento individual de cada jugador ya se encuentran disponibles, estas son utilizadas principalmente para entrenamientos, y no se encuentran disponibles a nivel comercial. Prevedemos que una segunda revolución de datos se originara una vez ese tipo de dato se haga disponible.

Adicionalmente nuestro modelamiento ignora el sistema de juego (formación) utilizada por los equipos o el entrenador de turno (con las implicaciones táctico-estratégicas que esto conlleva), y tampoco considera la demarcación funcional de cada jugador (posición dentro de la formación), ni tampoco el nivel de “presión” bajo el cual transcurre el encuentro (no es lo mismo jugar contra un equipo débil al comienzo del campeonato versus jugar contra el lider, peleando el campeonato, durante las ultimas fechas, en un equipo tradicionalmente fuerte). Sin duda, estos aspectos pueden condicionar las acciones que un jugador realiza durante el partido. Si bien estamos trabajando para incluir estos aspectos en un modelo, esto requiere un trabajo previo de etiquetación el rol de los jugadores en cada partido/formación, e identificación de una métrica razonable para medir la presión que enfrentan los jugadores. Aspectos adicionales que contribuyen potencialmente a mejorar nuestro modelo son, por ejemplo: i) hacer un manejo más detallado del transcurso del tiempo, incorporando heterogeneidad en la duración de las distintas acciones de los jugadores; ii) considerar como la disposición técnica de los entrenadores (sin necesariamente cambiar formación) afecta la toma de decisiones de los jugadores. Prevedemos que futuros esfuerzos de investigación se centraran en estos aspectos, y en la calibración del modelo utilizando otras ligas/temporadas.

Otros trabajos que se pueden realizar a partir de este es la inclusión de más ligas de fútbol profesional para ampliar el espectro de jugadores observables (mercado de futbolistas), utilizar otras métricas agregadas de impacto en los clubes tales como probabilidad de clasificar a copas internacionales y evitar el descenso, y por último, evaluar otras métricas del poder predictivo del modelo para mejorarlo y ser más precisos en este ámbito.

Agradecimientos: Esta investigación ha sido financiada en parte por el Instituto Sistemas Complejos de Ingeniería ISCI (ICM-FIC: P05-004-F, CO-

NICYT: FB0816).

Referencias

- [1] C. Anderson and D. Sally. *The Numbers Game: Why Everything you know about soccer is wrong*. Penguin Books, 2013.
- [2] Tom L. G. Bergkamp, Wouter G. P. Frencken, A. Susan M. Niessen, Rob R. Meijer, and Ruud. J. R. den Hartigh. How soccer scouts identify talented players. *European Journal of Sport Science*, 0(0):1–11, 2021. PMID: 33858300.
- [3] Dixon & Coles. Modelling association football scores and inefficiencies in the football betting market. *J Stor*, 0(0), 1997.
- [4] W. Dufour. Los métodos de objetivación del comportamiento motor en la recogida de datos en fútbol. *Apunts*, 1982.
- [5] FiveThirtyEight. Carmelo nba player predictions. accessed: 10/01/2022, 2018.
- [6] P. Galaz. Datazul: Un primer caso de analytics aplicado al fútbol profesional en Chile. accessed: 10/01/2022, 2020.
- [7] L. Gómez-Jordana, J. Milho, A. Ric, R. Silva, and P. Passos. Landscapes of passing opportunities in football – where they are and for how long are available? *Barça Sports Analytics Summit*, 2019.
- [8] S. Gregory. <https://www.statsperform.com/resource/expected-goals-in-context/>. accessed: 10/01/2022, 2020.
- [9] M. Hughes. Computerized notation analysis in field games. *Taylor and Frances*, 1988.
- [10] M. Lewis. *Moneyball: the art of winning an unfair game*. 2004.
- [11] G. Martín. Accumulated confirmed cases of coronavirus nationwide. accessed: 10/01/2022, 2020.
- [12] S. Mena. Impacto de los futbolistas de la premier league 2017-2018 en la probabilidad de salir campeón a través de un método de simulación con inferencia bayesiana. accessed: 01/05/2022, 2022.
- [13] R. Neal. Probabilistic inference using markov chain monte carlo methods. Technical report, University of Toronto, 1993.

- [14] E. Olsen and O. Larsen. *Science and Football III: Use of match analysis by coaches*. 1997.
- [15] L. Pappalardo, L. Guerrini, A. Rossi, and P. Cintia. Explainable injury forecasting in soccer via multivariate time series and convolutional neural networks. *Barça Sports Analytics Summit*, 2019.
- [16] L. Pappalardo and E. Massucco. Soccer match event dataset. *Figshare*, 2019.
- [17] D. Partridge and IM. Franks. A detailed analysis of crossing opportunities from the 1986 world cup (part i). *Soccer Journal*, 1989.
- [18] Á. Ric, R. Peláez, and Barça Innvotation HUB. *Football Analytics: Now and Beyond : a Deep Dive Into the Current State of Advanced Data Analytics*. 2019.
- [19] L. Shaw and M. Glickman. Dynamic analysis of team strategy in professional football. *Barca Sports Analytics Summit*, 2019.
- [20] B. Spencer, M. Hawkey, and M. Robertson. Using contextual player movement and spatial control to analyse player passing trends in football. *Barça Sports Analytics Summit*, 2019.
- [21] Stan Development Team. Stan modeling language users guide and reference manual. Version 2.29, 2022.
- [22] Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2 edition, 2009.
- [23] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

OPTIMIZACIÓN DE LA ZONIFICACIÓN Y EL RUTEO DE CAMIONES EN LA RECOLECCIÓN DE RESIDUOS DE UN MUNICIPIO DEL GRAN BUENOS AIRES

FEDERICO BERTERO *
MANUELA CERDEIRO †
GUILLERMO A. DURÁN ‡ §
NAZARENO A. FAILLACE ‡

Resumen

Este artículo presenta el uso de técnicas matemáticas y computacionales para la optimización del sistema de recolección de residuos urbanos de Berazategui, un municipio argentino del Gran Buenos Aires. Debido a que la organización manual no lograba cumplir con las expectativas de los vecinos y de los trabajadores, el objetivo no sólo es mejorar la eficiencia sino también lograr una distribución equitativa de trabajo entre recolectores y entre los conductores de los camiones. Debido a las características particulares del sistema de recolección de Berazategui y al tamaño del problema, el mismo fue particionado en 3 etapas. En la primera, una heurística construye zonas de recolección estructuralmente simples y balanceadas con respecto al número de cuadras que incluyen. En la segunda etapa, para cada una de las zonas generadas en la etapa inicial se diseñan rutas de longitud mínima para los camiones recolectores mediante un modelo de programación lineal entera mixta. La etapa final consistió en la asignación de camiones a las zonas de recolección generadas de modo de tener una asignación equilibrada con respecto a la duración de las jornadas laborales. Dado que son varios los factores que influyen en el tiempo de recolección, se optó por resolver este problema con un esquema de optimización biobjetivo, utilizando un modelo de

*Findo, Argentina.

†Departamento de Computación, FCEyN-UBA, Argentina.

‡Instituto de Cálculo FCEyN-UBA y CONICET, Argentina.

§Departamento de Ingeniería Industrial, Universidad de Chile, Chile.

programación lineal entera junto a un algoritmo iterativo. Los desarrollos aquí presentados fueron implementados por las autoridades de Berazategui a partir de 2020. En comparación con la organización manual que se encontraba en vigencia, las soluciones obtenidas con las herramientas desarrolladas logran una distribución de trabajo más de 70 % más equitativa y permiten ahorrarle a la municipalidad aproximadamente USD 34000 anuales en combustible.

Palabras Clave: Gestión de residuos sólidos urbanos, Diseño de rutas, Optimización, Programación lineal entera mixta.

1. Introducción

La gestión de residuos sólidos urbanos (RSU) se considera un desafío para las ciudades contemporáneas debido al rápido crecimiento en la cantidad de residuos, los altos costos de recolección de los mismos y las capacidades limitadas de tratamiento. De las actividades asociadas a los RSU, las de recolección y transporte representan el mayor porcentaje del costo total, por lo que mejoras en estos aspectos tienen como consecuencia inmediata un importante ahorro para los municipios [29].

La municipalidad de Berazategui está ubicada en la zona sur del Gran Buenos Aires, a 26 kilómetros de la ciudad de Buenos Aires, la capital del país. Tiene una superficie de 217km^2 con una población estimada al 2020 de 365.000 habitantes. A lo largo del tiempo, cada vez que la Municipalidad adquiría nuevos camiones, las zonas y las asignaciones de chóferes y recolectores eran modificadas a mano. Esto repercutía en una distribución muy desigual del trabajo y en el aumento de reclamos por parte de los vecinos insatisfechos con el servicio. Por otro lado, las rutas de los camiones no estaban preestablecidas, por lo que podría existir un margen de mejora en cuanto a la longitud de su recorrido diario y, en consecuencia, una reducción en el gasto de combustible y mantenimiento. Debido a la ineficiencia que ocasionaba la organización manual del procedimiento, las autoridades de la Municipalidad nos convocaron para mejorar el servicio de recolección de residuos. Específicamente, requerían el rediseño de la zonificación del municipio y la traza de rutas eficientes para los vehículos, procurando equidad de distribución de trabajo entre recolectores y entre choferes.

Por lo tanto, el objetivo de este artículo es el desarrollo e implementación de heurísticas y algoritmos basados en programación lineal entera para la mejora de la recolección y transporte de RSU de Berazategui. Se busca que las herramientas desarrolladas no sólo brinden soluciones de buena calidad

sino que también sean eficientes en cuanto a tiempo de ejecución, para poder ser fácilmente reutilizadas cuando la Municipalidad adquiera nuevos vehículos. Dada la complejidad asociada a los problemas donde se busca abordar la construcción de zonas de recolección, el diseño de rutas y la asignación de camiones a zonas [22] [28], no es posible obtener una solución óptima que los englobe en simultáneo. En nuestro caso, decidimos trabajar sobre cada uno de manera individual, buscando obtener la mejor solución que cumpliera con las restricciones propias del problema asociado y lo planteado por el ente del municipio encargado de realizar las tareas asociadas a la recolección de RSU.

Para diseñar las zonas de recolección implementamos una heurística que las construye iterativamente y luego intenta mejorar la forma de su frontera para que sean más sencillas de recordar para los choferes. Una vez definida la zonificación, se buscan rutas eficientes mediante un modelo de programación lineal entera mixta. Finalmente, teniendo en cuenta los resultados de las dos etapas anteriores, implementamos el Método de la Restricción Épsilon para lograr una repartición equilibrada de trabajo entre los choferes.

Este artículo está organizado de la siguiente manera. La siguiente sección consiste de una revisión de otros trabajos de optimización de gestión de RSU. En la tercera sección se brinda una descripción del problema, detallando cómo se realiza la recolección actualmente y de los requisitos planteados por el municipio de Berazategui. En la cuarta sección se describe cómo se recopilaron y procesaron los datos, y se presentan las técnicas matemáticas y computacionales desarrolladas para resolver cada uno de los tres problemas abordados. En la quinta sección se muestran los resultados numéricos que surgen de la implementación de cada uno de los esquemas de resolución. Finalmente, en la última sección se comentan las conclusiones y sugerencias para trabajo a futuro.

2. Revisión de la bibliografía

Uno de los primeros reportes bibliográficos sobre recolección residuos donde se exploran diferentes técnicas matemáticas para resolver esta clase de problemas fue realizada en 1974 por Beltrami y Bodin [3]. Desde entonces, dada la gran cantidad de potenciales aplicaciones y la versatilidad de éstas, numerosas investigaciones fueron realizadas y publicadas. En [18] los autores definen uno de los primeros algoritmos diseñados para generar simultáneamente la zonificación y el ruteo de vehículos para un sistema de recolección de residuos. Diversas aplicaciones de técnicas de optimización matemática sobre problemas similares de diversas ciudades del mundo aparecen en la literatura.

En [26] los autores llevan a cabo la optimización de rutas de recolección para una empresa estadounidense, que se traducen en importantes ahorros de logística. En [7] se aplican heurísticas de búsqueda local para definir los sectores de recolección y rutas eficientes para minimizar el tiempo de viaje de los vehículos y balancear la carga de trabajo. Por otro lado, en [20] se desarrolla un modelo de optimización multiobjetivo para determinar la ubicación de centros de recolección y la zonificación de áreas urbanas, y se aplican a dos casos de estudio en Irán. Otros ejemplos de aplicaciones incluyen a las ciudades de Chicago [12], Kaohsiung [6], Lisboa [21], Santiago de Chile [2], Turin [13], Ibarra [15], Karbala [27] y diferentes localidades argentinas [4], [5], [8] (en este último caso tratándose de un problema un poco diferente, dado que trata de la recolección de hojas en la ciudad estudiada). Finalmente, un importante estudio sobre problemas de ruteo de vehículos aplicados en la recolección de residuos se puede encontrar en [17].

Como se detalla en la siguiente sección, las características particulares de la recolección de Berazategui requieren la aplicación y el desarrollo de nuevas herramientas. Según lo revisado en la literatura, las tres etapas de nuestro enfoque de solución son contribuciones que pueden ser adaptadas a sistemas de recolección con procedimientos y disponibilidad de datos similares al caso de estudio.

3. Descripción del problema

En el municipio de Berazategui, los residuos domiciliarios se depositan en cestos particulares en la vereda y son recolectados por camiones que recorren un total de aproximadamente 15000 cuadras diarias, segmentadas en 42 zonas. El equipo de cada camión está formado por un chófer y 2 ó 3 recolectores de bolsas que las recogen y suben al camión. La jornada laboral está dividida en dos turnos. Los recolectores recorren una zona durante su jornada laboral mientras que, en general, cada chófer recorre dos zonas, una por la mañana y otra por la tarde, en el mismo camión. La carga de cada camión debe llevarse hacia la planta de tratamiento de residuos dependiente del organismo estatal denominado CEAMSE, ubicada a 23 km de distancia del Municipio. El traslado de la carga se realiza al finalizar la recolección, o al alcanzarse la capacidad máxima de carga del camión. Si la capacidad es alcanzada antes de finalizar la recolección, se deben hacer dos descargas en la planta del CEAMSE en una misma jornada, aumentando significativamente el horario laboral, el combustible requerido y el desgaste del vehículo.

Los objetivos planteados por la Dirección de Higiene Urbana (DHU) -el

ente del municipio encargado de realizar las tareas asociadas a la recolección de RSU- para la construcción de zonas de recolección fueron que las mismas sean uniformes respecto al metraje de recolección para que la carga laboral entre recolectores sea equitativa; que se respete la modalidad de dos zonas diarias para cada camión; que se evite en lo posible que un camión alcance la capacidad de carga antes de finalizar el recorrido; y, finalmente, conseguir una adjudicación pareja de pares de zonas entre los camiones para distribuir de manera justa el trabajo entre chóferes.

Por otro lado, para el diseño de rutas, la DHU solicitó que la ruta dentro de cada zona minimice el tiempo de recorrido y contemple la estrategia de *achique*, en la que el camión recorre una calle y los recolectores recorren las calles transversales juntando los residuos y acercándolos a la esquina por donde pasa el camión, donde otro recolector los toma y los sube al vehículo. Esta estrategia evita que el camión deba pasar por todas las cuadras, implicando una disminución del tiempo de recolección.

Las características especiales del problema a resolver hacen necesario el diseño de metodologías ad-hoc, diferentes a lo conocido en la literatura para problemas de recolección tal cual fue detallado en la sección anterior.

En primer lugar, el objetivo del trabajo no sólo se trata de encontrar una zonificación y diseñar rutas eficientes sino que también la municipalidad hizo especial hincapié en balancear la carga laboral entre los recolectores y entre los chóferes. En ambos casos fue necesario tener en cuenta los factores que influían en la duración de la jornada laboral de cada grupo. En el primero, el metraje de la zona asignada y, en el segundo, el peso total de residuos recolectados y el kilometraje de recorrido diarios. En el último caso, deben tenerse en cuenta dos objetivos independientes al momento de llevar a cabo la asignación de chóferes a zonas para garantizar una distribución equitativa de trabajo: minimizar la máxima disparidad entre la cantidad de residuos recolectados y minimizar la máxima diferencia entre longitud del recorrido. Esto implica un problema de optimización multiobjetivo.

En segundo lugar, el *achique* es la principal particularidad del sistema de recolección de Berazategui y es la razón por la cual el problema del diseño de rutas sea distinto a los encontrados en la literatura mencionada en la sección anterior. En este caso, no es necesario que el camión recorra todas cuadras de la zona, sino que debe garantizarse que visite al menos una vez todas las esquinas y que recorra las cuadras con una longitud mayor a 130 mts., las cuales resultan demasiado extensas para practicar el *achique*.

Al momento de iniciar este estudio, el Municipio tenía una zonificación, un ruteo de los camiones y una asignación de camiones a zonas realizada de manera manual, que a todas luces resultaba ineficiente, generando reclamos

de los vecinos por el mal servicio y protestas de los trabajadores del sector por el trabajo desigual asignado a cada uno de ellos.

4. Enfoque de solución

Para llevar adelante el proyecto fue necesario primero recopilar y procesar todos los datos vinculados con la recolección de residuos del municipio. La información reunida la siguiente:

- Un mapa del municipio indicando sentidos de calles, ubicación de semáforos, vías de tren y sus cruces, arroyos, obstáculos o problemas de tránsito.
- Datos de las zonas y rutas de recolección utilizadas hasta el momento.
- Las cantidades (en kilaje) de los residuos recolectados diariamente por cada camión recolector, obtenidas a través de los registros de las descargas hechas en la planta del CEAMSE.
- Datos demográficos geolocalizados.
- Datos de GPS de los recorridos diarios de los camiones.

Se implementó una rutina que examina la ruta de cada camión obtenida a partir del GPS y reconoce, día por día, qué zonas recolectó y cuánta distancia recorrió. Esta información luego fue procesada junto con la cantidad de residuos descargados y con los datos demográficos para lograr una aproximación de la cantidad de residuos diarios que genera cada manzana del municipio, lo cual fue utilizado como dato en la resolución del problema. En otras, para procurar que el trabajo de los chóferes sea lo más equitativo posible y evitar que deban descargar en la planta del CEAMSE más de una vez por día.

También se mantuvieron entrevistas con chóferes y supervisores para adquirir un panorama más general del problema y entender qué factores influían en la longitud de sus jornadas laborales.

Una vez realizada la recopilación y el procesamiento de todos los datos necesarios, el problema a resolver fue particionado en tres etapas debido a su tamaño y complejidad computacional: diseño de zonas de recolección, traza de rutas eficientes y asignación de camiones a zonas.

4.1. Diseño de zonas

Tanto en el diseño de zonas como en la elaboración de rutas de los camiones recolectores, se utilizó el plano de la Municipalidad de Berazategui alojado en

la base de datos de OpenStreetMap (OSM) [23]. El plano es esencialmente un archivo en formato `xml` que consta, entre otros, de dos tipos de objetos: *nodo* (o *node*) que representa un punto con coordenadas de latitud y longitud y *camino* (o *way*) que es una línea que conecta nodos. Para este trabajo, se empleó la herramienta `osmfilter`¹, para filtrar en el plano todos aquellos caminos en los que no pudiera acceder el camión recolector y aquellos que se encontraran fuera de los límites de la Municipalidad. Para visualizar el plano se utilizó Java OpenStreetMap Editor² (JOSM). En lo que resta de este análisis, denominaremos esquina a cada nodo del plano y calle a la línea que conecta dos esquinas.

Nuestro modelo de zonificación utiliza las manzanas como unidades elementales a partir de las cuales se construyen las zonas de recolección. Es decir, cada zona de recolección es un conjunto de manzanas. Consideramos como *manzanas* a los espacios del plano delimitados por calles. Además, requerimos que cumpla las siguientes características. En primer lugar, una manzana no puede contener a otras manzanas. Por otro lado, una manzana no debe ser atravesada por vías de ferrocarril ni por carriles de una autopista; de esta manera, se busca evitar demoras en el recorrido debido a desvíos en la ruta del camión recolector. El reconocimiento de las manzanas se llevó a cabo mediante una sencilla rutina que permite identificarlas en la abstracción del plano del municipio, a través del recorrido en ambos sentidos de cada una de sus cuadras.

Diremos que dos manzanas son adyacentes si tienen en común al menos una de las calles de sus fronteras. Por otro lado, dado un conjunto A de manzanas, definimos su metraje de recolección como la suma de la longitud de sus calles, en metros, y lo notaremos $m(A)$. La longitud de cada una de las calles es sumada una sola vez aunque delimite dos manzanas del conjunto.

Una vez identificadas las manzanas, se elaboró un grafo $G = (V, E)$ cuyos vértices representan a las manzanas y donde dos vértices están conectados por una arista si y sólo si las manzanas correspondientes son adyacentes. Se notará con B_v a la manzana correspondiente al vértice $v \in V$ y con v_B al vértice correspondiente a la manzana B . En la 1 se presenta un ejemplo de la relación biunívoca entre las manzanas de un plano y los vértices de su grafo. Un conjunto de manzanas A se dice conexo si y sólo si el subgrafo inducido en G por los vértices correspondientes a las manzanas de A es conexo.

El proceso de zonificación del municipio estuvo guiado por dos características deseables de las zonas que la componen: balance respecto al metraje de recolección y cuadratura de su frontera. La primera busca asegurar relativa

¹<https://wiki.openstreetmap.org/wiki/Osmfilter> (Accedido en 2019)

²<https://josm.openstreetmap.de/wiki> (Accedido en 2019)

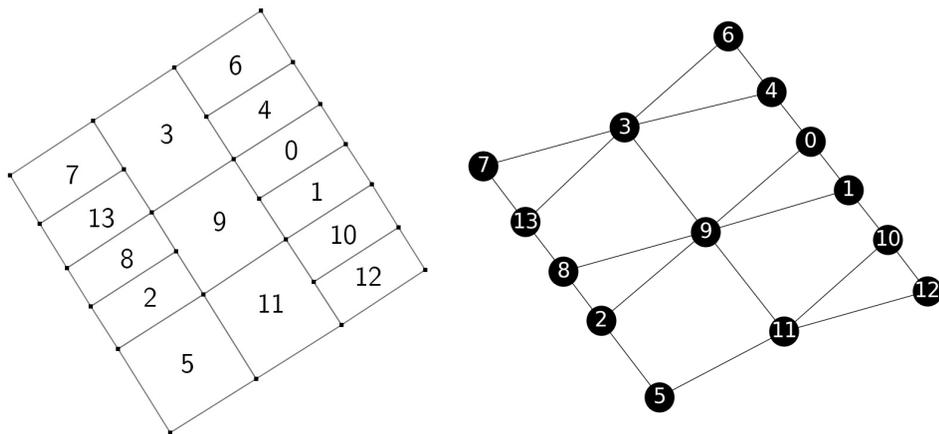


Figura 1: Una sección del plano con sus manzanas y calles y su correspondiente grafo G .

equidad en la duración de las jornadas laborales y el esfuerzo físico que deben realizar los recolectores, mientras que la segunda pretende simplificar la memorización de los límites de las zonas para los chóferes. Por otra parte, la única restricción consistía en que cada zona fuera conexa. Para asegurar su cumplimiento, el proceso de zonificación se llevó a cabo en cada una de las secciones del plano que determinaban las componentes conexas G_s de G . A cada una de esas secciones se le calculó el metraje de recolección y, con el asesoramiento de Higiene Urbana, se determinó en cuántas zonas se dividiría, teniendo en cuenta, entre otros factores, la densidad poblacional y la ubicación de edificios públicos y comercios. De esta manera, para cada sección s , elaborar una nueva zonificación consiste en definir n_s zonas de recolección, de manera tal que cada una tenga metraje de recolección de alrededor de $\frac{m(s)}{n_s}$ metros y que su frontera sea tan parecida a un cuadrado como sea posible. En otras palabras, dejando de lado por un momento la cuadratura de las zonas, para cada G_s se busca una partición balanceada en n_s bloques. Dado que se ha demostrado que no hay un algoritmo polinomial para este tipo de problemas (salvo que $P = NP$) [1], y considerando además que se busca que una zona sea lo más cuadrada posible, se decidió desarrollar una heurística que permita obtener un resultado en un período de tiempo razonable y lo suficientemente bueno para ser aplicado en el municipio.

Antes de describir el procedimiento de la heurística, se presenta el criterio que se utilizó para medir cuán parecida es la frontera de una zona a un cuadrado. Para definirlo, se consideraron dos tipos de medidas. La primera de ellas fue la medida de cuadratura desarrollada en [25], en el contexto de procesamiento de imágenes. Sea $A \subset \mathbb{R}^2$ acotado y tal que ∂A sea una curva cerrada

y simple, los autores exponen una medida \mathcal{Q} para determinar cuán parecida es ∂A a un cuadrado. Esta medida cumple las siguientes propiedades:

- $\mathcal{Q}(A) \in (0, 1]$ para todo A .
- $\mathcal{Q}(A) = 1 \iff \partial A$ es un cuadrado.
- $\mathcal{Q}(A)$ es invariante respecto a traslaciones, rotaciones y escalas.

La medida \mathcal{Q} viene dada por la siguiente fórmula:

$$\mathcal{Q}(A) = \frac{\sqrt{2}}{3} \frac{\text{area}(A)}{\min_{\alpha \in [0, 2\pi]} \iint_{A(\alpha)} (|x| + |y|) dx dy} \quad (1)$$

donde $A(\alpha)$ denota la rotación de A en un ángulo α alrededor de su centroide. La otra medida considerada fue la medida de convexidad planteada en [8] la cual, en el contexto de recolección de hojas, fue desarrollada con el objetivo de formar conjuntos de manzanas cuadradas adyacentes cuya frontera se pareciera lo más posible a un cuadrilátero. Sea $A \subset \mathbb{R}^2$ y sea $\text{conv}(A)$ su cápsula convexa, la medida de convexidad \mathcal{M} viene dada por:

$$\mathcal{M}(A) = \frac{\text{area}(A)}{\text{area}(\text{conv}(A))} \quad (2)$$

Es fácil ver que $\mathcal{M}(A) \in (0, 1]$ y que \mathcal{M} es invariante por traslación, rotación y escalamiento. Pensando en un plano ideal correspondiente a una grilla perfecta, $\mathcal{M}(A)$ toma valores muy cercanos a 1 si ∂A describe un cuadrilátero.

Ambas medidas son aplicables a una zona cuya frontera define una curva cerrada y simple. Sin embargo, en pruebas experimentales, ninguna de las dos medidas por sí misma brindó resultados completamente satisfactorios. Por un lado, \mathcal{Q} prácticamente ignora pequeñas perturbaciones en la frontera de una zona con decenas de manzanas. Por otro lado, \mathcal{M} otorga valores muy altos a zonas cuya frontera se asemeja a un rectángulo alargado. Por esta razón, se decidió utilizar como criterio de medida de cuadratura de una zona a una combinación convexa entre \mathcal{Q} y \mathcal{M} , que notaremos \mathcal{S} . Sea $A \subset \mathbb{R}^2$ tal que pueda aplicársele \mathcal{Q} , se define \mathcal{S} como :

$$\mathcal{S}(A) = \lambda \mathcal{Q}(A) + (1 - \lambda) \mathcal{M}(A) \quad \lambda \in (0, 1) \quad (3)$$

Notar que si ∂A tiene la forma de un cuadrado, entonces $\mathcal{S}(A) = 1$. Tras diversas pruebas experimentales, $\lambda = 0,75$ ha resultado el parámetro más adecuado para realizar el proceso de zonificación en el plano de Berazategui.

Con la intención de simplificar la notación en la descripción de la heurística desarrollada, notaremos $G = (V, E)$ en vez de G_s y n en vez de n_s . Como se detallará más adelante, en cada iteración del procedimiento se definirá al menos una zona de recolección, por lo tanto notaremos como $G^{(k)} = (V^{(k)}, E^{(k)})$ al subgrafo de G inducido por los vértices correspondientes a manzanas del plano que no forman parte de ninguna zona luego de la k -ésima iteración del procedimiento. Así, se define $G^{(0)} = G$. Recordemos que en las manzanas de G se deseaba definir n zonas de recolección. Además, dado $F = (V_F, E_F)$ un subgrafo inducido de G , se notará $\mathcal{B}(F) = \{B_v\}_{v \in V_F}$. Se define el peso del grafo $G^{(k)}$ como $m(G^{(k)}) = m(\{B: v_B \in G^{(k)}\})$. Se indicará con $\mathcal{A}^{(k)}(Z)$ al conjunto de manzanas adyacentes a las manzanas de la zona Z en la sección del plano correspondiente a $G^{(k)}$. Se utilizará m_{MIN} para notar el mínimo metraje de recolección deseable que puede tener una zona. Por último, \mathcal{Z} denotará el conjunto de zonas de recolección ya definidas en el plano.

A continuación, se describe el procedimiento de la heurística. Se inicializan $k = 0$ y $\mathcal{Z} = \emptyset$, y se itera hasta que $|\mathcal{Z}| = n$. Al comienzo de la k -ésima iteración, si $|\mathcal{Z}| = n - 1$, se define una nueva zona Z con las manzanas de $\mathcal{B}(G^{(k)})$, se la añade a \mathcal{Z} y culmina el procedimiento. De lo contrario, se calcula $\bar{v} = \arg \max_{v \in V^{(k)}} \varepsilon(v)$, donde $\varepsilon(v)$ denota la excentricidad de v en $G^{(k)}$. Interpretándolo geográficamente, $B_{\bar{v}}$ es una manzana que se encuentra en alguno de los bordes del área del plano que todavía no ha sido zonificada. Luego, se inicializa una nueva zona $Z = \{B_{\bar{v}}\}$. Mientras $|\mathcal{Z}| < \frac{m(G^{(k)})}{n - |\mathcal{Z}|}$, iterativamente se añade a Z la manzana $\tilde{B} = \arg \max_{B \in \mathcal{A}^{(k)}(Z)} \mathcal{S}(Z \cup \{B\})$, es decir, la manzana que maximiza la cuadratura de la zona, dentro de todas las manzanas adyacentes que aún no son parte de ninguna zona. Una vez que $|\mathcal{Z}| \geq \frac{m(G^{(k)})}{n - |\mathcal{Z}|}$, se incorpora Z a \mathcal{Z} . Luego se define $H^{(k)}$ como el grafo obtenido a partir de eliminar de $G^{(k)}$ los vértices correspondientes a las manzanas de Z . Si $H^{(k)}$ es conexo, se define $G^{(k+1)} = H^{(k)}$ y se continúa con la $(k + 1)$ -ésima iteración.

De no ser así, se nota $H_1^{(k)}, \dots, H_J^{(k)}$ ($J > 1$) a las componentes conexas de $H^{(k)}$ en orden descendiente de metraje. Diremos que $H_j^{(k)}$ es grande si $m(H_j^{(k)}) \geq m_{MIN}$ o si $j = 1$ (pues a esta altura del procedimiento falta definir por lo menos una zona de recolección y hace falta una iteración más); si no, diremos que es pequeña. Como se detallará en seguida, en las componentes grandes se definirán nuevas zonas de recolección y las manzanas de las pequeñas se repartirán entre las zonas ya demarcadas que limiten con ellas. En primer lugar, se define $G^{(k+1)} = H_1^{(k)}$. Para cada $H_j^{(k)}$ grande se pueden definir a lo sumo $\lfloor \frac{m(H_j^{(k)})}{m_{MIN}} \rfloor$ zonas de recolección. Podría no ser trivial decidir cuántas zonas delimitar en cada una de ellas pues hay que observar que cada zona tendrá deseablemente un metraje igual al promedio del metraje de las

zonas ya definidas. Los conjuntos de manzanas de las componentes grandes que no sean asignadas a ninguna zona se repartirán entre las zonas limítrofes, por lo tanto es deseable que sean lo más pequeños posible. Para lograr esto, se planteó un modelo de Programación Lineal Entera Mixta (MILP) que permita decidir cuántas zonas se formarán en cada una de las componentes conexas grandes (se lo denota como MILP1). En el modelo se introduce como restricción que se elabore al menos una zona en $H_1^{(k)}$, dado que es la componente conexas con mayor metraje y con la que se trabajará en la $(k + 1)$ -ésima iteración. Se trata de un modelo muy sencillo cuya formulación se encuentra en la subsección Modelos auxiliares. Seguido a la resolución del modelo, se delimitan las nuevas zonas en cada componente conexas grande tal que $j \neq 1$ invocando a la heurística de manera recursiva sobre el área del plano que le corresponde, introduciendo \mathcal{Z} como zonificación inicial y la solución del modelo de MILP como cantidad de zonas a definir. Si en alguna componente conexas grande el modelo indicara no formar zonas de recolección, se la pasa a considerar como una componente pequeña.

Para cada componente conexas pequeña, las manzanas de $\mathcal{B}(H_j^{(k)})$ son repartidas entre las q zonas que las rodean. Se ejecuta recursivamente la heurística sobre el sector del plano que le corresponde para definir q subzonas de recolección con las manzanas de $\mathcal{B}(H_j^{(k)})$, y luego se utiliza un modelo de MILP, que se denota como MILP2 y cuya formulación se encuentra en la subsección Modelos auxiliares, para decidir a cuál de las q zonas lindantes se anexa cada una de las subzonas, garantizando su conexión. Si alguna de las subzonas no limita con ninguna de las zonas, se aplica recursivamente la heurística y el modelo de MILP para repartir sus manzanas entre las subzonas vecinas. El modelo de MILP utilizado tiene como objetivo minimizar el valor absoluto de la diferencia de metraje entre las q zonas luego del anexo de las subzonas. Seguidamente, se procede a la $(k + 1)$ -ésima iteración. En 1 se muestra el pseudocódigo del cuerpo principal de la heurística, mientras que en 2 se encuentra el pseudocódigo del procedimiento que se lleva a cabo con las componentes conexas de $H^{(k)}$.

Es importante notar que el procedimiento no puede incurrir en recursión infinita pues la cantidad de manzanas en el plano es finita. Por otro lado, al comienzo de cada iteración, la conexión de $G^{(k)}$ está garantizada. El procedimiento de la heurística puede aplicarse con otros criterios de balance de zonificación. Por ejemplo, si se quisiera lograr equilibrio respecto a la cantidad de residuos en cada zona en vez de su metraje de recolección, basta considerar la cantidad estimada de residuos generada por cada manzana en vez de su metraje.

Puesto que el criterio de parada de la generación de una zona está relacio-

Algoritmo 1: ALGORITMO ZONIFICADOR

Entrada: G, \mathcal{Z}, n

Salida: \mathcal{Z}

mientras $|\mathcal{Z}| < n$ **hacer**

si $|\mathcal{Z}| = n - 1$ **entonces**

$Z \leftarrow \mathcal{B}(G)$

$\mathcal{Z} \leftarrow \mathcal{Z} \cup \{Z\}$

en otro caso

$\bar{v} \leftarrow \arg \max_{v \in V} \varepsilon(v)$

$Z \leftarrow \{B_{\bar{v}}\}$

mientras $|\mathcal{Z}| < \frac{m(G)}{n-|\mathcal{Z}|}$ **hacer**

$\tilde{B} \leftarrow \arg \max_{B \in \mathcal{A}(Z)} \mathcal{S}(Z \cup \{B\})$

$Z \leftarrow Z \cup \{\tilde{B}\}$

$\mathcal{Z} \leftarrow \mathcal{Z} \cup \{Z\}$

$H \leftarrow$ subgrafo inducido en G por $V - \bigcup_{B \in \mathcal{Z}} v_B$

si H es conexo **entonces**

$G \leftarrow H$

en otro caso

$G, \mathcal{Z} \leftarrow$ SUBRUTINA_COMPONENTES_CONEXAS(H, \mathcal{Z}, n)

devolver \mathcal{Z}

Algoritmo 2: SUBRUTINA_COMPONENTES_CONEXAS

Entrada: H, \mathcal{Z}, n

Salida: G, \mathcal{Z}

$CCH \leftarrow$ componentes conexas de H ordenadas decrecientemente según m

$G \leftarrow H_1$

$H_{BIG} \leftarrow \{H_j \in CCH : m(H_j) \geq m_{MIN}\}$

$H_{SMALL} \leftarrow CCH - H_{BIG}$

si $n - |\mathcal{Z}| = 1$ **entonces**

$H_{SMALL} \leftarrow H_{SMALL} \cup (H_{BIG} - \{H_1\})$

en otro caso

$S \leftarrow$ solución de MILP1 (S_j : número de zonas a definir en $\mathcal{B}(H_j)$)

para cada $H_j \in H_{BIG} : j \neq 1$ **hacer**

si $S_j > 0$ **entonces**

$\mathcal{Z} \leftarrow$ ALGORITMO ZONIFICADOR($H_j, \mathcal{Z}, |\mathcal{Z}| + S_j$)

en otro caso

$H_{SMALL} \leftarrow H_{SMALL} \cup \{H_j\}$

para cada $H_j \in H_{SMALL}$ **hacer**

$A \leftarrow$ zonas adyacentes a $\mathcal{B}(H_j)$ in \mathcal{Z}

$subzones \leftarrow$ ALGORITMO ZONIFICADOR($H_j, \emptyset, |A|$)

 resolver MILP2

 aneXar cada subzona a la zona en A indicada por la solución de MILP2

devolver G, \mathcal{Z}

nado con el metraje, puede ocurrir que la frontera sea susceptible de mejoras removiendo o agregando algunas manzanas. Por lo cual, se implementó un algoritmo de corrección que se aplica a la zonificación obtenida con la heurística. Se trata de un algoritmo de búsqueda local basado en Steepest Ascent Hill Climbing [19] y su objetivo es minimizar la cantidad de vértices del perímetro de cada zona. Sean s y t dos calles que forman parte de la frontera de la zona Z tales que tienen una esquina en común c , y sea $\alpha \in (-\pi, \pi]$ el ángulo entre s y t , diremos que Z tiene un vértice en c si y sólo si $|\alpha| \in (\frac{1}{5}\pi, \frac{4}{5}\pi)$. De esta manera, la cantidad de vértices de Z viene dada por la cantidad de esquinas de su frontera que son vértices y se nota $v(Z)$.

Por otro lado, sea \mathcal{B} el conjunto de manzanas del plano, diremos que cierta manzana $b \in Z$ es una manzana fronteriza de Z si existe alguna manzana del plano $\hat{b} \in \mathcal{B} - Z$ tal que b es adyacente a \hat{b} . En ese caso diremos que \hat{b} es una manzana adyacente a Z . Notaremos con ∂Z al conjunto de manzanas fronterizas de Z y con $\mathcal{A}(Z)$ a las manzanas adyacentes a Z . En cada iteración del algoritmo, a cada zona se le anexa la manzana adyacente o se le quita la manzana fronteriza que logre la mayor reducción en el número de vértices, si la hubiere. Recordar que la zonificación \mathcal{Z} es una partición de \mathcal{B} , por lo tanto, cada manzana pertenece a exactamente una zona y anexar o quitar una manzana provoca un cambio en alguna de las zonas aledañas a Z . Como la transformación que sufre cada zona impacta en su metraje, es prudente establecer un límite para cuánto puede crecer o encogerse una zona con respecto a su estado inicial. Para esto, se definen $r \geq 1$ un radio aceptable de modificación y antes de modificar las zonas, para cada Z , el siguiente conjunto:

$$M_Z = \{\hat{b} \in \mathcal{B} - Z : \min_{b \in Z} \{d(v_{\hat{b}}, v_b)\} \leq r\} \cup \{\hat{b} \in Z : \min_{b \in \partial Z} \{d(v_{\hat{b}}, v_b)\} \leq r - 1\} \quad (4)$$

donde d es la distancia en el grafo G y v_b es el vértice de G que le corresponde a $b \in \mathcal{B}$. En la 2 se muestra un ejemplo de M_Z a partir del estado inicial de una zona Z . Vale aclarar que M_Z se mantiene fijo durante todo el procedimiento, mientras que $\mathcal{A}(Z)$ y ∂Z se redefinen cada vez que Z es modificada. A lo largo del algoritmo, a Z sólo pueden anexarse las manzanas que se encuentren en $M_Z \cap \mathcal{A}(Z)$ y sólo se le pueden quitar las manzanas en $M_Z \cap \partial Z$. Para cada $b \in \mathcal{B}$ se nota con $W(b)$ a la zona a la que pertenece b .

Otro aspecto de la zonificación que desea mantenerse es el de la conexión de todas sus zonas. Por esto, antes de anexar o quitar una manzana, se verifica que el cambio no provoque la desconexión de alguna de las zonas involucradas. Además, para evitar deshacer un cambio que se llevó a cabo en otra zona durante la misma iteración, se utiliza una lista tabú con longitud igual a la cantidad de zonas donde se lleva el registro de los cambios hechos, para no deshacerlos. El algoritmo finaliza cuando se alcanza un límite de iteraciones

o cuando no se ha producido ningún cambio durante una iteración. En 3 se presenta el pseudocódigo del algoritmo corrector, donde (X, Y, b) nota la transferencia de la manzana b desde la zona X a la zona Y .

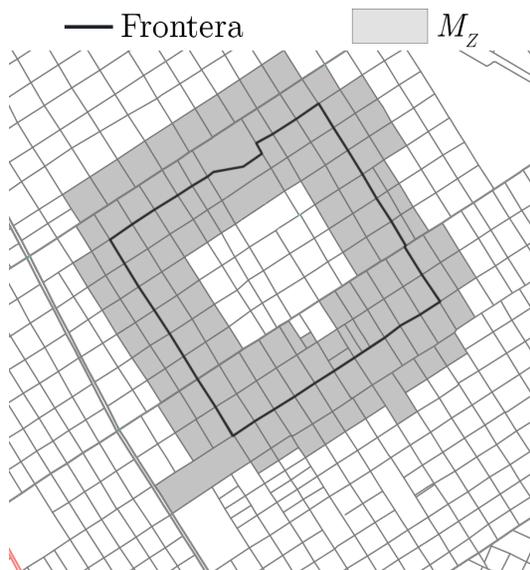


Figura 2: Ejemplo del conjunto M_Z para una zona Z , con $r = 2$.

4.1.1. Modelos auxiliares

A continuación se presentan los modelos de Programación Lineal Entera auxiliares empleados en el diseño de las zonas. En esta sección nos referiremos a la medida de un conjunto de manzanas como su metraje o como el peso de residuos generados por las manzanas que la componen. Cualquiera de los dos criterios se pueden utilizar para trabajar con las componentes conexas.

El primer modelo, denominado MILP1, permite decidir cuántas zonas de recolección se crearán en cada una de las componentes conexas grandes de $H^{(k)}$, $\{H_i^{(k)}\}_{i \in I}$. Los parámetros del modelo son los siguientes:

- I : el conjunto de índices de las componentes conexas grandes de $H^{(k)}$.
- c : la cantidad de zonas que faltan ser creadas.
- m : la medida mínima que debe tener una zona de recolección.
- M : la medida deseable que debe tener una zona de recolección.
- w_i : la medida de $H_i^{(k)}$ para $i \in I$.

Algoritmo 3: ALGORITMO CORRECTOR

Entrada: $\mathcal{Z}, \mathcal{B}, r, iter_max$ **Salida:** \mathcal{Z} **para cada** $Z \in \mathcal{Z}$ **hacer**| Construir M_Z $T \leftarrow$ inicializar lista tabú de longitud $|\mathcal{Z}|$ $it \leftarrow 0$ **mientras** $it \leq iter_max$ **y al menos alguna zona cambia** **hacer**| **para cada** $Z \in \mathcal{Z}$ **hacer**| | $V \leftarrow v(Z)$ | | $best_change \leftarrow \emptyset$ | | **para cada** $b \in M_Z \cap \mathcal{A}(Z): W(b) - \{b\}$ **es conexa** **hacer**| | | $Z \leftarrow Z \cup \{b\}$ | | | **si** $v(Z) < V$ **y** $(Z, W(b), b) \notin T$ **entonces**| | | | $V \leftarrow v(Z)$ | | | | $Z \leftarrow Z - \{b\}$ | | | | $best_change \leftarrow (W(b), Z, b)$ | | **para cada** $b \in M_Z \cap \partial Z: Z - \{b\}$ **es conexa** **hacer**| | | $Z \leftarrow Z - \{b\}$ | | | **si** $v(Z) < V$ **y** $(W(b), Z, b) \notin T$ **entonces**| | | | $V \leftarrow v(Z)$ | | | | $Z \leftarrow Z \cup \{b\}$ | | | | $best_change \leftarrow (Z, W(b), b)$ | | **si** $best_change \neq \emptyset$ **entonces**| | | Actualizar T agregando $best_change$

| | | Realizar la transferencia de manzanas indicada por

| | | | $best_change$ | | $it \leftarrow it + 1$ **devolver** \mathcal{Z}

Para cada $i \in I$, se definen una variable entera x_i que indica la cantidad de zonas de recolección que serán creadas en $H_i^{(k)}$ y una variable continua z_i que cuantifica la medida del sector sin zonificar en $H_i^{(k)}$ que resultaría de generar x_i zonas de recolección con medida M .

$$\text{mín} \quad \sum_{i \in I} z_i \quad (5)$$

$$\text{s.t.:} \quad \sum_{i \in I} x_i = c \quad (6)$$

$$x_i \leq \frac{w_i}{m} \quad \forall i \in I \quad (7)$$

$$w_i - Mx_i \leq z_i \quad \forall i \in I \quad (8)$$

$$x_1 \geq 1 \quad (9)$$

$$x_i \in \mathbb{Z}_{\geq 0} \quad \forall i \in I$$

La función objetivo (5) minimiza la medida total de los sectores que podrían no quedar zonificados y que deberían ser repartidos entre las zonas lindantes. Con respecto a las restricciones, (6) asegura que se crearán la cantidad de zonas de recolección necesarias; (7) acota superiormente la cantidad de zonas que se pueden crear en cada $H_i^{(k)}$ según su medida; (8) establece la relación entre las variables y (9) asegura que en $H_1^{(k)}$ se genere al menos una zona de recolección.

El segundo modelo auxiliar, MILP2, es utilizado para repartir las subzonas entre las zonas limítrofes. Se supone que cada subzona limita con al menos una zona ya definida. De lo contrario, se le aplica la heurística de zonificación y se reparten las sub-subzonas entre las subzonas vecinas con este mismo modelo. Se definen los siguientes parámetros:

- I : conjunto de subzonas.
- J : conjunto de zonas de recolección entre las que se repartirán las subzonas.
- δ_{ij} : vale 1 si y sólo si la subzona $i \in I$ limita con la zona $j \in J$.
- w_i : medida de la subzona $i \in I$.
- W_j : medida de la zona $j \in J$ sin anexar ninguna subzona.

Para cada par $(i, j) \in I \times J$ se define una variable binaria x_{ij} que indica si la subzona i es anexada a la zona j . Además, para cada $j \in J$ se introduce una variable continua y_j que cuantifica la medida de j luego de llevar a cabo la asignación. Por último, se emplea una variable continua z para modelar el máximo valor absoluto de la diferencia de medida post-asignación entre todos los pares zonas de J , es decir, $z = \max_{(j,k) \in J^2} |y_j - y_k|$.

$$\text{mín} \quad z \quad (10)$$

$$\text{s.t.} \quad x_{ij} \leq \delta_{ij} \quad \forall i \in I \forall j \in J \quad (11)$$

$$\sum_{j \in J} x_{ij} = 1 \quad \forall i \in I \quad (12)$$

$$\sum_{i \in I} w_i x_{ij} + W_j = y_j \quad \forall j \in J \quad (13)$$

$$y_j - y_k \leq z \quad \forall (j, k) \in J^2 \quad (14)$$

$$x_{ij} \in \{0, 1\} \quad \forall i \in I \forall j \in J$$

La función objetivo (10) minimiza la disparidad entre la zona con mayor medida y la de menor medida. Las restricciones (11) previenen que una subzona sea anexada a una zona no limítrofe; (12) aseguran que cada subzona debe ser anexada a exactamente una zona; (13) relacionan a las variables de asignación con las variables de medida y (14) establecen el valor absoluto de las diferencias de medidas entre todos los pares de zonas en J como cota inferior para z .

4.2. Diseño de rutas

Una vez definidas las zonas de recolección, el siguiente paso consistió en buscar las rutas de longitud mínima que deberían recorrer los camiones recolectores. Recordemos que en la Municipalidad es frecuente que la recolección se lleve a cabo con un procedimiento que se conoce como achique (). Por esta razón, es obligatorio que el vehículo recorra las cuadras largas y que transite todas las esquinas de la zona. A diferencia de otros problemas de recolección de residuos, no es necesario que el camión recorra todas las cuadras de la zona.

El plano de la municipalidad puede representarse en un grafo mixto $P = (V_P, A_P, E_P)$ donde A_P es el conjunto de arcos que representan a las calles de una mano y E_P el conjunto de aristas que representan a las calles doble mano.

A continuación se elabora el digrafo $G = (V, A)$ donde $V = V_P$, $(i, j) \in A \forall (i, j) \in A_P$ y para cada $(i, j) \in E_P$, los arcos (i, j) y (j, i) están en A . Cada arco (i, j) de G tiene como peso a la longitud de la calle que corresponde al arco o arista de P y se nota $\ell(i, j)$. Por otro lado, el plano contiene información sobre las restricciones de tránsito. Estas son procesadas como caminos en G y almacenadas en el conjunto F . Por ejemplo, si el recorrido $(i, j) \rightarrow (j, k)$ no está permitido, $(i, j, k) \in F$. Los caminos asociados a las restricciones tienen tres o cuatro nodos.

Como suele ocurrir en los problemas de ruteo, se requiere conocer el camino mínimo entre nodos. Para esto se utilizó el algoritmo desarrollado en [10], que

permite hallar el camino mínimo (no necesariamente simple) desde un nodo fuente al resto de los nodos del grafo sin transitar completamente por ninguno de los caminos prohibidos de F . Se referirá a este algoritmo como dPPG y , si $(i, j) \notin A$, se notará con γ_{ij} al camino mínimo entre i, j en G que no tiene ningún camino prohibido como subcamino. Si $(i, j) \in A$, consideramos $\gamma_{ij} = (i, j)$. Dado un camino $\gamma = (i_0, i_1, \dots, i_N)$, notaremos como su longitud en metros a $\ell(\gamma) = \sum_{k=0}^{N-1} \ell(i_k, i_{k+1})$.

De aquí en adelante se nota con $s \in V$ al nodo correspondiente al punto de partida de los camiones recolectores y con $t \in V$ al punto de llegada. Estos nodos corresponden a dos esquinas distintas del predio de Higiene Urbana. En el diseño de rutas para los camiones recolectores resulta esencial conocer el camino mínimo en G desde s a v para cada $v \in V - \{s\}$ y desde cada $v \in V - \{t\}$ hasta t . Para lo primero, se aplicó dPPG a G con s como nodo fuente. Para lo segundo, se invirtieron los caminos obtenidos de aplicar dPPG al digrafo traspuesto de G con t como nodo fuente.

A continuación, sea Z una zona de recolección, consideramos $V_Z \subset V$ al conjunto de nodos que son esquinas de alguna de las manzanas de Z . Finalmente se define el digrafo completo $T = (V_T, A_T)$ con $V_T = V_Z \cup \{s, t\}$ y el arco (i, j) tiene como peso w_{ij} a $\ell(\gamma_{ij})$. Como el camión debe recorrer todas las esquinas de Z partiendo desde s y llegando a t , entonces se trata de resolver una instancia del Asymmetrical Travelling Salesman Problem (ATSP) en T , con restricciones adicionales. Para esto se utilizó un modelo de Programación Lineal Entera Mixta basado en el modelo propuesto por Desrochers y Laporte [9] para resolver el ATSP. Si bien existen modelos con relajaciones lineales más ajustadas [24], el de Desrochers y Laporte fue el que mejor se desempeñó para nuestra aplicación.

Para el planteo del modelo se definen los siguientes parámetros:

- L^{MAX} : longitud máxima de una cuadra que puede ser achicada.
- $\mathcal{L} = \{e \in A_S \cup E_S : \ell(e) > L^{MAX}\}$ el conjunto de cuadras que no pueden ser achicadas.
- L_e : para cada $e \in \mathcal{L}$, $(i, j) \in A_T$ pertenece a L_e si en γ_{ij} se recorre alguno de los arcos correspondientes a e .
- V_{ij} : para $i, j \in V_T$, $i \neq j$ se define V_{ij} como el conjunto de $k \in V_T$ tales que, si u es el anteúltimo nodo de γ_{ij} y v es el segundo de γ_{jk} , $(u, j, v) \in F$.
- F_4 : caminos prohibidos de cuatro nodos.

- B_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a B_δ si y sólo si el último arco de γ_{vu_2} es (u_1, u_2) .
- E_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a E_δ si y sólo si el primer arco de γ_{u_3v} es (u_3, u_4) .
- B_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a B_δ si y sólo si el último arco en γ_{vu_2} es (u_1, u_2) .
- \hat{B}_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a \hat{B}_δ si y sólo si los últimos dos arcos γ_{vu_3} son (u_1, u_2) y (u_2, u_3) .
- E_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a E_δ si y sólo si el primer arco en γ_{u_3v} es (u_3, u_4) .
- \hat{E}_δ : para cada $\delta = (u_1, u_2, u_3, u_4) \in F_4$, $v \in V_T$ pertenece a \hat{E}_δ si y sólo si los primeros dos arcos en γ_{u_2v} son (u_2, u_3) y (u_3, u_4) .

Para cada arco $(i, j) \in A_T$ se define una variable binaria x_{ij} que vale 1 si y sólo si γ_{ij} forma parte del recorrido. Además, para cada $i \in V_T$ se define una variable real no negativa u_i que indica la posición del nodo i en el recorrido del grafo T .

$$\text{mín} \quad \sum_{i \in V_T} \sum_{\substack{j \in V_T \\ j \neq i}} \ell(\gamma_{ij}) x_{ij} \quad (15)$$

$$\text{s.a:} \quad \sum_{i \in V_T} x_{ij} = 1 \quad \forall j \in V_T \quad (16)$$

$$\sum_{j \in V_T} x_{ij} = 1 \quad \forall i \in V_T \quad (17)$$

$$u_i - u_j + (n-1)x_{ij} + (n-3)x_{ji} \leq n-2 \quad \forall i, j \in V_T - \{s\} \quad (18)$$

$$-u_i + (n-3)x_{is} + \sum_{\substack{j \in V_T - \{s\} \\ j \neq i}} x_{ji} \leq n-1 \quad \forall i \in V_T - \{s\} \quad (19)$$

$$u_i + (n-3)x_{si} + \sum_{\substack{j \in V_T - \{s\} \\ j \neq i}} x_{ji} \leq n-1 \quad \forall i \in V_T - \{s\} \quad (20)$$

$$\sum_{(i,j) \in L_e} x_{ij} \geq 1 \quad \forall e \in \mathcal{L} \quad (21)$$

$$x_{ts} = 1 \quad (22)$$

$$\sum_{k \in V_{ij}} x_{jk} \leq 1 - x_{ij} \quad \forall i, j \in V_T \quad (23)$$

$$\sum_{v \in B_\delta} x_{vu_2} + x_{u_2u_3} + \sum_{v \in E_\delta} x_{u_3v} \leq 2 \quad \forall \delta \in F_4 \quad (24)$$

$$\sum_{v \in \hat{B}_\delta} x_{vu_3} + \sum_{v \in E_\delta} x_{u_3v} \leq 1 \quad \forall \delta \in F_4 \quad (25)$$

$$\sum_{v \in B_\delta} x_{vu_2} + \sum_{v \in \hat{E}_\delta} x_{u_2v} \leq 1 \quad \forall \delta \in F_4 \quad (26)$$

$$u_i \geq 0 \quad \forall i \in V_T$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in A_T$$

El objetivo (15) es minimizar la longitud del recorrido que visita exactamente una vez los nodos de T . Las restricciones (16)-(20) corresponden al modelo de ATSP con nodo de partida s , de Desrochers y Laporte: (16) y (17) garantizan que cada nodo de T es visitado exactamente una vez, mientras que (18)-(20) evitan la formación de subtours. Por otro lado, (21) asegura que el camión recorrerá todas las cuadras que son demasiado largas como para ser achicadas. (22) determina que t es el último nodo visitado. Por construcción,

γ_{ij} no tiene ningún camino prohibido. Sin embargo, entre la finalización de un camino mínimo y el comienzo del siguiente podría violarse una restricción de tránsito. (23)-(26) impiden esta situación al contemplar los casos de caminos prohibidos de tres y cuatro nodos.

A partir de la solución del modelo, la ruta con la que se recorre Z partiendo y llegando al predio de Higiene Urbana se elabora uniendo de manera ordenada los caminos mínimos γ_{ij} tales que $x_{ij} = 1$.

4.3. Asignación de camiones a zonas

La última etapa del trabajo consistió en asignar cada uno de los camiones a las zonas de recolección, con el objetivo de equilibrar la duración de la jornada laboral de sus chóferes. Diariamente cada camión es conducido por un chófer y la única característica que diferencia a los camiones es su capacidad de carga. A cada camión le corresponden dos zonas y la secuencia diaria suele ser la siguiente: el camión parte de Higiene Urbana con el equipo de recolectores del turno mañana, recorre la primera zona, regresa al predio de Higiene Urbana para cambiar el equipo de recolectores, recorre la segunda zona, visita Higiene Urbana para dejar a los recolectores, viaja a CEAMSE a descargar los residuos y finalmente el camión es regresado a Higiene Urbana. Los últimos dos pasos de la secuencia son comunes a todos a los camiones recolectores y, por lo tanto, no son tenidos en cuenta al momento de calcular la longitud total de su recorrido. Se considera que la longitud del recorrido de una zona viene dada por la longitud de la ruta obtenida en la sección anterior.

Por otro lado, la cantidad de residuos generados por las zonas asignadas a un camión también tienen impacto en la duración de la jornada laboral de su chófer. Una zona podría tener una ruta no muy extensa, pero si tiene considerable densidad poblacional, genera más desechos y se consume más tiempo para recolectarlos. Además, el exceso de la capacidad máxima del camión supone un viaje extra a CEAMSE y una extensión de la jornada laboral de varias horas para el chófer y los recolectores.

Se identificaron dos características de las zonas de recolección que influyen en la duración de la jornada laboral de los chóferes: la distancia total de sus recorridos y la cantidad de residuos que generan. Sin embargo, no es sencillo determinar cuantitativamente cuánto influye cada una de ellas. Por esa razón, se enfocó este problema desde la perspectiva de la optimización biobjetivo. Como suele suceder en problemas de optimización combinatoria multiobjetivo, es posible que no exista una solución que cumpla ambos objetivos simultáneamente [11]. Por lo tanto, se trata de hallar un conjunto de soluciones Pareto-óptimas.

Para aproximar el Frente de Pareto se empleó el Método de la Restricción Épsilon, que consiste en resolver el problema de optimización considerando sólo uno de los objetivos e incluir el resto como restricciones del problema [14]. En el caso de la optimización biobjetivo de minimización, se busca $\min_{x \in X} \{f_1(x) | f_2(x) \leq \varepsilon\}$; modificando sistemáticamente el valor de ε se puede hallar el Frente de Pareto [16]. Se comienza resolviendo el problema con un único objetivo $\min_{x \in X} f_1(x)$. Sea x^1 su solución, a continuación se resuelve el problema $\min_{x \in X} \{f_1(x) | f_2(x) \leq f_2(x^1) - \delta\}$, y así sucesivamente, hasta que el problema resulte infactible. Finalmente, se resuelve $\min_{x \in X} f_2(x)$.

Seguidamente se presenta la formulación del modelo de asignación de camiones a zonas como un modelo de Programación Lineal Entera Mixta, que se utilizó para aplicar el Método de la Restricción Épsilon. Los parámetros del problema son los siguientes:

- T : el conjunto de camiones recolectores
- Z : el conjunto de zonas
- c_i : la capacidad de carga del camión $i \in T$
- w_j : el peso de los residuos generados por la zona $j \in Z$
- ℓ_j : la longitud del recorrido de la zona $j \in Z$

Para cada $i \in T$, $j \in Z$ se define la variable binaria x_{ij} , que indica si la zona j es asignada al camión i .

$$\text{mín} \quad y \quad (27)$$

$$\text{s.t.} \quad \sum_{j \in Z} x_{ij} = 2 \quad \forall i \in T \quad (28)$$

$$\sum_{i \in T} x_{ij} = 1 \quad \forall j \in Z \quad (29)$$

$$\sum_{j \in Z} x_{ij} w_j \leq c_i \quad \forall i \in T \quad (30)$$

$$\sum_{j \in Z} (x_{ij} - x_{kj}) w_j \leq y \quad \forall i, k \in T \quad (31)$$

$$\sum_{j \in Z} (x_{ij} - x_{kj}) \ell_j \leq \varepsilon \quad \forall i, k \in T \quad (32)$$

$$x \in \{0, 1\}^{n \times m}$$

$$y \in \mathbb{R}$$

La función objetivo (27) y la restricción (31) modelan la minimización de la máxima disparidad de carga. Con respecto a las demás restricciones del problema, (28) asegura que a cada camión sean asignadas exactamente dos zonas, (29) garantiza que a cada zona le sea asignado exactamente un camión, (30) que la capacidad de ningún camión sea excedida y (32) es la restricción ϵ que involucra al objetivo de minimizar la máxima disparidad en distancia recorrida.

5. Resultados

En esta sección mostramos los resultados obtenidos en el diseño del sistema de recolección de la Municipalidad. La implementación de los modelos de las tres etapas (zonificación, ruteo y asignación de pares de zonas a camiones) se llevó a cabo en Python 3.7 y los modelos de Programación Lineal Entera se resolvieron utilizando CPLEX 12.8 y su API para Python. Los tiempos de ejecución corresponden a una computadora con procesador Intel Core i5 @ 2.60GHz con 16 GB de RAM.

Los resultados son comparados con la situación previa a la finalización de nuestro trabajo, utilizando los datos proporcionados por la Municipalidad mencionados al comienzo de la sección . Más precisamente se comparan la zonificación obtenida con la que utilizaba la Municipalidad en marzo de 2020; las distancias promedio recorridas por los camiones en nuestro resultado con las distancias promedio según los datos de GPS desde julio de 2019 hasta febrero de 2020 (excluyendo en ambos casos los viajes a CEAMSE al final de la jornada) y la asignación de chóferes a zonas propuesto por nuestro enfoque de solución con el que estaba en vigencia en marzo de 2020.

5.1. Zonificación

Desde Higiene Urbana nos indicaron que deseaban mantener algunas de las zonas con las que contaban y definir manualmente otras. Esto se debía a que en ciertos lugares de la ciudad la recolección de residuos tenía una organización especial que no les resultaba conveniente modificar: el barrio El Pato, el barrio Juan María Gutiérrez, los barrios cerrados y algunos sectores céntricos. Omitiendo dichas zonas, restaba zonificar un área de alrededor de 47 km² con un metraje a cubrir de 680 km.

Se aplicó la heurística descrita en la 4.1 a cada una de sus componentes conexas, imponiendo que cada zona debía recolectar al menos 3150 kg. de residuos. En la 3a se puede observar el plano del Municipio y las zonas generadas

por la heurística, mientras que en la 3b se muestra la zonificación obtenida luego de aplicarle el algoritmo corrector. Las partes sombreadas corresponden a las secciones de la ciudad que no requerían ser rezonificadas. En la 1 se muestra la comparación de distintos aspectos del metraje entre la zonificación en vigencia antes de realizar el proyecto y las obtenidas con los métodos que desarrollamos.

Como se puede apreciar, la zonificación previa a nuestro trabajo presentaba una marcada disparidad en el metraje a cubrir de las zonas de recolección, que se traducía en una distribución de trabajo muy desigual entre los recolectores. Por otro lado, el algoritmo corrector no perjudicó demasiado la diferencia en el metraje entre las zonas y permitió que sus fronteras fueran más sencillas de recordar para los chóferes. Además de proponer una zonificación que representara una distribución de trabajo más equitativa para los recolectores, se logró evitar que haya zonas no conexas: en la zonificación previa seis zonas de recolección tenían dos o más componentes conexas, mientras que en las propuestas obtenidas con nuestros dos métodos todas las zonas son conexas.

	Manual	Heurística	Corregido
Desvío estándar	7,84 km	1,82 km	1,89 km
Max	46,15 km	26,62 km	27,13 km
Min	5,03 km	20,53 km	20,34 km

Tabla 1: Comparación del metraje de la zonificación que utilizaba la Municipalidad antes del proyecto con las propuestas desarrolladas en este trabajo.

Desde Higiene Urbana se mostraron muy satisfechos con los resultados obtenidos. Se implementaron algunas modificaciones puntuales con el objetivo de mejorar su cuadratura y considerando aspectos demográficos. De esta manera, se obtuvo la nueva zonificación del municipio y se dio inicio al diseño de rutas para los camiones.

5.2. Diseño de rutas

El diseño de rutas se llevó a cabo para todas las zonas salvo para las correspondientes al barrio El Pato y a los barrios cerrados, que cuentan con un esquema de recolección distinto. A través de un análisis descriptivo del territorio y de consultas con Higiene Urbana, se concluyó que cualquier cuadra con longitud mayor a 130 metros sería considerada larga y, por ende, el camión debería recorrerla obligatoriamente.

Para cada una de las zonas, se resolvió el modelo de la 4.2, con las siguientes especificaciones. Si a los 45 minutos el solver no logró encontrar una solución



(a)



(b)

Figura 3: (a) Zonificación resultante de aplicar la heurística a las áreas del plano que se deseaban rezonificar. (b) Zonificación resultante de aplicar el algoritmo corrector a la zonificación de (a). El área sombreada corresponde a sectores de la ciudad que no debían ser zonificados.

óptima se le permite correr otros 45 minutos para hallar una solución con un gap de optimalidad menor a 1 %. En caso de no encontrarla, se permite que continúe con la resolución del modelo por una hora y media más, configurando CPLEX de modo de mejorar la cota inferior del valor óptimo y el gap de optimalidad. En diversas experimentaciones se observó que luego de la primera hora y media de búsqueda de soluciones, en general no se hallaban nuevas soluciones factibles con mejor valor en la función objetivo, por lo que se decidió darle más énfasis a mejorar la cota.

De las 38 zonas para las que se diseñó la ruta de recolección, en 20 se hallaron rutas óptimas con un tiempo promedio de 11 minutos y 25 segundos. Por otro lado, para 17 zonas se hallaron rutas subóptimas con gap de optimalidad de a lo sumo 1 %, en un tiempo promedio de 1 hora y 27 minutos. En estos casos, el gap de optimalidad promedio fue de 0,7 % y el gap absoluto promedio fue de 172,54 metros. La zona correspondiente al barrio José María Gutiérrez requirió un tiempo más prolongado de resolución, dado que se trata de una zona con 308 nodos (más del doble de cantidad de nodos que el promedio de las demás). Por esta razón, se duplicaron los límites de tiempo para el solver. En 2 horas y 44 minutos se obtuvo una solución con un gap de optimalidad del 3,31 % y un gap absoluto de 1,61 km.

Previo a nuestro trabajo, el promedio de kilómetros diarios recorridos por cada camión eran 62,87 km. mientras que con las nuevas rutas propuestas esa cantidad descendió a 48,63 km. De esta manera, en promedio se redujo en un 22,6 % la cantidad de kilómetros recorridos. Considerando que las nuevas rutas se diseñaron para 20 camiones, esta diferencia se traduce en una reducción total de aproximadamente 284,79 km de recorrido diario. Teniendo en cuenta que la municipalidad paga USD 0,98 por litro de combustible y cada uno rinde aproximadamente 3,5 km., la reducción en la distancia de recorrido se traduce en un ahorro anual de aproximadamente USD 34000. A ese monto hay que sumarle el ahorro en mantenimiento debido a la reducción del desgaste, lo cual es difícil de cuantificar pero no representa un gasto despreciable.

La ruta diseñada para cada zona fue presentada de tres maneras: un archivo `html` con la animación de la ruta sobre el plano del municipio, que permite una visualización sencilla y rápida; una hoja de ruta tradicional con la secuencia de indicaciones de giro y un archivo `gpx` para ser utilizado en la aplicación OsmAnd, que proporciona una guía GPS incluso si el usuario se encuentra offline.

5.3. Asignación de camiones a zonas

A partir del cálculo aproximado de la cantidad de residuos que genera cada zona y la longitud del recorrido que le corresponde, se asignaron dos zonas a cada camión, de manera tal que la carga laboral de los choferes sea lo más balanceada posible. Se aplicó el Método de la Restricción Épsilon tomando como objetivo principal minimizar la diferencia del peso que debía ser recolectado por cada camión e incluyendo la diferencia en kilómetros de recorrido diario como la restricción épsilon. En cada iteración, se disminuía el valor de este parámetro en 250 metros., hasta que el problema resultó infactible. Finalmente, se resolvió el problema considerando como único objetivo a la diferencia en kilómetros de recorrido diario. De esta manera, se obtuvo una aproximación del Frente de Pareto para el problema de optimización biobjetivo.

Cada una de las soluciones pareto-óptimas representaba una posible asignación de camiones a zonas. En total se lograron obtener 24 propuestas de asignaciones. En 4 se muestra la aproximación del Frente de Pareto y la asignación manual vigente antes de nuestro trabajo. Es importante destacar que se excluyen los camiones encargados de la recolección en José María Gutiérrez, El Pato, los barrios cerrados y la zona céntrica, dado que por solicitud de la Municipalidad esas asignaciones no debían modificarse. En la figura también se destaca la asignación definitiva elegida por el municipio. La decisión final de cuál elegir fue producto de la percepción y el conocimiento del problema de parte de los responsables operativos de la recolección.

En la situación previa a nuestro trabajo, la diferencia máxima de peso recolectado entre los camiones era de aproximadamente 6070 kg. y la disparidad máxima de kilómetros recorridos era de aproximadamente 44,5 km. En la asignación de nuestra propuesta elegida por la Municipalidad, dichas diferencias son de 1825 kg. y 7,72 km., respectivamente. Uno de los factores que contribuyó a mejorar el balance de trabajo entre los chóferes fue el rediseño de las zonas de recolección que permitió disminuir las zonas no conexas. Por otro lado, el proceso de asignación permitió corregir la situación en la que ambas zonas de algunos chóferes estaban geográficamente cerca de Higiene Urbana mientras que las de otros se encontraban más alejadas.

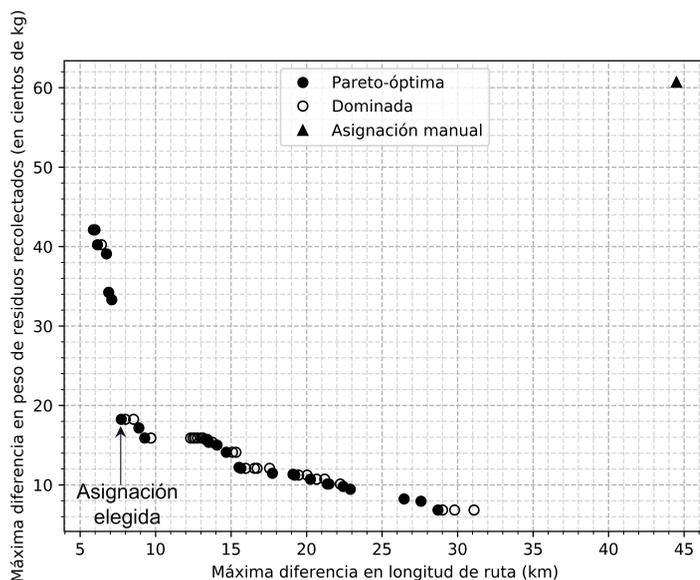


Figura 4: Aproximación del Frente de Pareto para el problema de optimización biobjetivo y comparación con la situación previa a nuestro trabajo.

6. Conclusiones

Se ha desarrollado una estrategia que permite mejorar el sistema de recolección de residuos del municipio de Berazategui. La implementación de una heurística de zonificación ha permitido reducir considerablemente la disparidad en la cantidad de trabajo entre los recolectores, mientras que la aplicación de modelos de programación entera ha posibilitado diseñar rutas eficientes y organizar el procedimiento del achique. Las rutas propuestas logran reducir en un 22,6% la distancia recorrida por los camiones, lo que trae un ahorro anual en combustible que ronda los USD 34,000 y una importante disminución del deterioro de los camiones recolectores. La asignación equitativa del par de zonas a cada camión permite balancear fuertemente el desgaste de los camiones y el trabajo de los conductores.

Pensando en la adaptabilidad de las herramientas diseñadas, cambiando algunos parámetros la estrategia propuesta en este trabajo podrá utilizarse para rediseñar las zonas y rutas de recolección si la municipalidad adquiere nuevos vehículos ante aumento de la población o mayor generación de residuos. Más aún, si bien los modelos fueron diseñados teniendo en cuenta las

características particulares de Berazategui, se pueden adaptar a servicios de recolección de otras ciudades, que tengan características similares en cuanto a su organización. La heurística de zonificación es teóricamente aplicable a cualquier plano mientras que el modelo de diseño de rutas puede aplicarse a sistemas de recolección en los que no sea necesario que el camión recorra todas las cuadras.

Con respecto al trabajo futuro para mejorar nuestra metodología, se puede continuar desarrollando la medida de cuadratura de una zona para que sea más sensible en el procedimiento iterativo de zonificación, de manera tal que el algoritmo corrector no deba realizar demasiadas modificaciones posteriormente. Además, se podría buscar un mejor parámetro para limitar el crecimiento de las zonas que se adapte mejor a mapas en los que el tamaño de las manzanas es muy dispar. Utilizando el promedio de metraje como dicho parámetro causó que las primeras zonas generadas fueran moderadamente más grandes que las últimas.

En lo que respecta al diseño de las rutas, es muy probable que el desempeño del modelo de programación entera pueda mejorarse en cuanto al tiempo necesario para hallar una solución óptima. Sin embargo, teóricamente la capacidad de los vehículos y la duración de la jornada laboral de los recolectores imponen restricciones al tamaño de las zonas que debe recorrer, por lo cual se espera no sea muy superior al de las analizadas en este trabajo.

Para concluir, el intendente del Municipio, Juan Patricio Mussi, mostró su satisfacción con el proyecto desarrollado. “Previo a este proyecto veíamos que agregábamos más camiones para la recolección pero no teníamos claro como aprovecharlos para mejorar la eficiencia del servicio. Los algoritmos de programación matemática implementados nos dieron la solución al problema: conseguimos un sistema de recolección que mejora y balancea las condiciones laborales de nuestros operarios, permite ahorrar combustible y desgaste de nuestros camiones y brinda un mejor servicio a los ciudadanos de Berazategui. Siempre he sido un defensor de la incorporación de herramientas científico-tecnológicas para la mejora de la gestión pública. Este proyecto nos mostró que vamos por el camino correcto. Seguramente podremos a futuro aplicar estas técnicas para mejorar el funcionamiento de otras áreas de nuestro municipio”, manifestó entusiasmado.

Agradecimientos: Este trabajo forma parte de un convenio de cooperación entre el Instituto de Cálculo de la Facultad de Ciencias Exactas y Naturales de la UBA y la municipalidad de Berazategui para asesorar a este municipio en la reformulación de su recolección de residuos urbanos. Agradecemos la permanente colaboración para concretar este proyecto de las autoridades

de Berazategui, en particular a Juan Patricio Mussi, Intendente de Berazategui durante la concreción del proyecto, a Diego Jevos, responsable político del sistema de recolección de residuos y a Ezequiel Lachowicz, quienes fueron fuente de consulta permanente durante todo el desarrollo del trabajo. Agradecemos también a Javier Marengo, por sus aportes para el desarrollo de las herramientas implementadas en este trabajo; y a Juan Doldan y Flavia Bonomo, por el vínculo inicial con las autoridades del municipio. El tercer autor está parcialmente financiado por el Instituto Sistemas Complejos de Ingeniería, ISCI, Chile (ICM-FIC: P05-004-F, CONICYT: FB0816), el subsidio UBACyT 20020170100495BA (Argentina) y el subsidio PIP-CONICET 11220200100084CO (Argentina).

Referencias

- [1] K. Andreev and H. Racke. Balanced graph partitioning. *Theory of Computing Systems*, 39:929–939, 2006.
- [2] C.A. Arribas, C.A. Blazquez, and A. Lamas. Urban solid waste collection system using mathematical modelling and tools of geographic information systems. *Waste Management & Research*, 28(4):355–363, 2010.
- [3] E. Beltrami and L. Bodin. Networks and vehicle routing for municipal waste collection. *Networks*, 4:65–94, 1974.
- [4] F. Bonomo, G. Duran, F. Larumbe, and J. Marengo. A method for optimizing waste collection using mathematical programming: a Buenos Aires case study. *Waste management & research : the journal of the International Solid Wastes and Public Cleansing Association, ISWA*, 30:311–24, 2011.
- [5] G. Braier, G. Duran, J. Marengo, and F. Wesner. An integer programming approach to a real-world recyclable waste collection problem in argentina. *Waste Management & Research*, 35(5):525–533, 2017.
- [6] N. Chang and H.Y. Lu. A new approach for long term planning of solid waste management systems using fuzzy global criterion. *Journal of Environmental Science and Health . Part A: Environmental Science and Engineering and Toxicology*, 32(4):1025–1047, 1997.
- [7] M. Cortinhal, M. Mourão, and A. Nunes. Local search heuristics for sectoring routing in a household waste collection context. *European Journal of Operational Research*, 255, 04 2016.

- [8] D. Delle Donne, V. Di Tomaso, and G. Durán. Optimizing leaf sweeping and collection in the argentine city of trenque lauquen. *Waste Management & Research*, 39(2):209–220, 2021.
- [9] M. Desrochers and G. Laporte. Improvements and extensions to the miller-tucker-zemlin subtour elimination constraints. *Operations Research Letters*, 10(1):27–36, 1991.
- [10] L. Di Puglia Pugliese and F. Guerriero. Dynamic programming approaches to solve the shortest path problem with forbidden paths. *Optimization Methods and Software*, 28(2):221–255, 2013.
- [11] M. Ehrgott and X. Gandibleux. *Multiobjective Combinatorial Optimization — Theory, Methodology, and Applications*, pages 369–444. Springer US, Boston, MA, 2002.
- [12] D. Eisenstein and A. Iyer. Garbage collection in chicago: A dynamic scheduling model. *Management Science*, 43:922–933, 1997.
- [13] E. Fadda, L. Gobbato, G. Perboli, M. Rosano, and R. Tadei. Waste collection in urban areas: A case study. *INFORMS Journal on Applied Analytics*, 4(48):307–322, 2018.
- [14] Y. Y. Haimes, L. Lasdon, and D. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-1(3):296–297, 1971.
- [15] Israel D. Herrera-Granda, Juan C. León-Jácome, Leandro L. Lorente-Leyva, Fausto Lucano-Chávez, Yackleem Montero-Santos, Winston G. Oviedo-Pantoja, and Christian S. Díaz-Cajas. Subregion districting to optimize the municipal solid waste collection network: A case study. In *Information and Communication Technologies of Ecuador (TIC.EC)*, pages 225–237. Springer International Publishing, 2019.
- [16] C.L. Hwang, A.S.M. Masud, S.R. Paidy, and K.S.P. Yoon. *Multiple Objective Decision Making, Methods and Applications: A State-of-the-art Survey*. Lecture notes in economics and mathematical systems. Springer-Verlag, 1979.
- [17] B. Kim, S. Kim, and S. Sahoo. Waste collection vehicle routing problem with time windows. *Computers & Operations Research*, 33(12):3624–3642, 2006. Part Special Issue: Recent Algorithmic Advances for Arc Routing Problems.

- [18] J.W. Male and J.C. Liebman. Districting and routing for solid waste collection. *Journal of the Environmental Engineering Division*, 104(1):1–14, 1978.
- [19] Z. Michalewicz. *How to Solve It: Modern Heuristics 2e*. Springer-Verlag, Berlin, Heidelberg, 2010.
- [20] S. Mostafayi Darmian, S. Moazzeni, and L. Magnus Hvattum. Multi-objective sustainable location-districting for the collection of municipal solid waste: Two case studies. *Computers & Industrial Engineering*, 150, 2020.
- [21] M.C. Mourão and M.T. Almeida. Lower-bounding and heuristic methods for a refuse collection vehicle routing problem. *European Journal of Operational Research*, 121(2):420–434, 2000.
- [22] R. Nauss. Solving the generalized assignment problem: An optimizing and heuristic approach. *INFORMS Journal on Computing*, 15:249–266, 2003.
- [23] OpenStreetMap contributors. Datos obtenidos de . <https://www.openstreetmap.org>, Accedido en 2019. Página web operativa en 2022.
- [24] R. Roberti and P. Toth. Models and algorithms for the asymmetric traveling salesman problem: an experimental comparison. *Journal of Mechanical Systems for Transportation and Logistics*, 1:113–133, 2012.
- [25] P. Rosin and J. Žunić. Measuring squareness and orientation of shapes. *Journal of Mathematical Imaging and Vision*, 39:13–27, 2011.
- [26] S. Sahoo, S. Kim, B. Kim, B. Kraas, and A. Popov. Routing optimization for waste management. *INFORMS Journal on Applied Analytics* 35, 1:24–36, 2005.
- [27] F. Salah Sahib and N. Shakir Hadi. Truck route optimization in karbala city for solid waste collection. *Materials Today: Proceedings*, 2021.
- [28] M. Solomon. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Oper. Res.*, 35:254–265, 1987.
- [29] E. B. Tirkolaee, P. Abbasian, M. Soltani, and S. A. Ghaffarian. Developing an applied algorithm for multi-trip vehicle routing problem with time windows in urban waste collection: A case study. *Waste Management & Research*, 37(1_suppl):4–13, 2019. PMID: 30761957.

REMUNERACIÓN DE REDES DE TRANSMISIÓN EN CHILE POR VNR: PRINCIPIOS, PROBLEMAS Y RECOMENDACIONES

RONALD FISCHER ^{*†}
RODRIGO MORENO ^{‡§†}

Resumen

Este trabajo analiza la regulación de las redes de transmisión en Chile. Se concentra en particular en el uso del mecanismo de Valor Nuevo de Reemplazo o VNR para remunerar las redes existentes. Se muestra, usando los principios de regulación moderna, que el VNR no es el mecanismo apropiado. El problema es que, a diferencia del pasado, actualmente el regulador determina las inversiones a realizar en el sector. Por lo tanto, es innecesario introducir riesgos en la remuneración de activos para dar señales de inversión eficiente. El trabajo muestra la magnitud de este efecto considerando la diferencia en el margen sobre costos referenciales de subastas de obras nuevas y de obras de ampliación, ambas con un menor nivel de riesgo que el VNR. El trabajo realiza un análisis de las distintas fuentes de riesgo para un inversionista en transmisión bajo VNR, distinguiendo los factores que favorecen a las empresas de aquellos que tienden a perjudicarla. Finalmente, se realiza una propuesta de un mecanismo regulatorio que no tiene los problemas del VNR en el caso de instalaciones preexistentes de transmisión.

Palabras Clave: Redes de transmisión, Inversión eficiente, Modelos regulatorios.

*Departamento de Ingeniería Industrial, Universidad de Chile, Chile.

†Instituto Sistemas Complejos de Ingeniería (ISCI), Chile.

‡Departamento Ingeniería Eléctrica, Universidad de Chile, Chile.

§Imperial College London, Reino Unido.

1. Introducción

Un sistema de regulación o de control tarifario requiere incentivos a la inversión. Esto significa que el sistema debe estar diseñado para atraer inversiones necesarias para responder a los requerimientos de la sociedad. El objetivo de la regulación eficiente es hacerlo al menor costo posible para los usuarios. Existen distintos enfoques regulatorios que intentan resolver este problema, desde los sistemas cost-plus hasta enfoques modernos que enfatizan los incentivos. Entre ellos, se encuentra el enfoque CPI-X (ahora modificado a *RIIO* : $Revenue = Incentives + Innovation + Outputs$) del Reino Unido de los 90, o el de empresa modelo chilena utilizada en distribución eléctrica, en telefonía (aunque ahora aplicado solo a los servicios de terminación de red), y en el sector sanitario (con algunos ajustes). Independientemente de sus ventajas teóricas, la calidad de los mecanismos regulatorios depende de la forma de su implementación y uso en el mundo real.

El objetivo de este trabajo es estudiar el mecanismo regulatorio de valorización de las instalaciones de transmisión, tanto desde el punto de vista teórico como en su aplicación. El foco particular de este trabajo es la remuneración de las instalaciones existentes de transmisión, las cuales son valorizadas siguiendo el método del Valor Nuevo de Reemplazo (VNR). Nuestra evaluación muestra que la actual aplicación del VNR a las instalaciones existentes no es apropiada, por lo que se proponen modificaciones normativas como solución de largo plazo, así como propuestas para el corto plazo.

Las recomendaciones normativas que se derivan de este trabajo consisten en utilizar los resultados de las licitaciones de obras de ampliación como una referencia para fijar los VI de las instalaciones existentes, reduciendo así el riesgo que introducen las revalorizaciones aleatorias del VI en el actual modelo. Segundo, se propone reducir la remuneración de aquellas obras que han cumplido su vida útil regulatoria, para reducir la sobrerenta que generan.

Es importante aclarar que, desde un punto de vista conceptual, el proceso tarifario contiene dos aspectos. El primero se centra en cómo se determina la remuneración necesaria para que las empresas inviertan y operen en forma eficiente; y el segundo se focaliza en cómo se recaudan los ingresos desde los usuarios de la red. En estos dos aspectos es posible introducir incentivos a la eficiencia. En este trabajo, no obstante, solo se examina en detalle el primer aspecto, es decir, la remuneración a las inversiones de las empresas.

La estructura de este documento es la siguiente. La Sección 2 describe los

conceptos fundamentales asociados a la remuneración de redes eléctricas y que se utilizan en este trabajo. La Sección 3 describe y analiza en profundidad el mecanismo de VNR, focalizándose en las condiciones conceptuales para que opere adecuadamente, así como los riesgos observados en la práctica. En la Sección 4 se lleva a cabo una discusión que finalmente justifica las recomendaciones de modificaciones regulatorias del estudio.

2. Marco teórico

En esta sección, se describen los conceptos fundamentales asociados a la remuneración de redes eléctricas y que se utilizarán, en las siguientes secciones, para analizar el marco regulatorio chileno y para identificar potenciales problemas y soluciones.

2.1. Modelos regulatorios¹

Existen principalmente dos esquemas básicos para regular las tarifas en sistemas regulados. Primero, están aquellos esquemas basados en costos reales y, segundo, aquellos basados en incentivos. En la literatura, estos se denominan mecanismos de bajo y alto poder, respectivamente. Un mecanismo de bajo poder (puro) es uno en que -en principio- la remuneración está asociada a los costos efectivos/reales de la empresa. Esto significa que la empresa enfrenta bajo riesgo, porque no tendría pérdidas. Por lo mismo, el retorno mínimo requerido para llevar a cabo las inversiones que requiere el sistema es bajo. El inconveniente de este mecanismo es que no existen incentivos a reducir costos, ni tampoco a hay incentivos a realizar inversiones de forma eficiente (el llamado efecto Averch-Johnson [3]).

Por el contrario, en un mecanismo de alto poder (puro), la remuneración de la empresa no está asociada a sus costos reales, sino a los que tendría una empresa eficiente -el benchmark- en cada periodo tarifario, como ocurre en el caso de la distribución en Chile. Por lo tanto, la empresa enfrenta fuertes incentivos a reducir sus costos. Esto también produce que la empresa enfrente un mayor nivel de riesgo, porque puede ser que el benchmark (e.g., empresa modelo) tenga costos menores (mayores) a los costos de la empresa real². Por

¹Ver [12]. El análisis de esta sección no considera la posibilidad de expropiación regulatoria, porque ella hace insostenible a la empresa regulada.

²Más adelante se discute cómo el esquema de empresa modelo basado en VNR puede resultar en valores mayores o menores a los costos reales de las empresas, incluso si las empresas se comportan de forma eficiente.

lo tanto, ante esta potencial varianza en los retornos, el costo de capital exigido por el inversionista es mayor. Otra consecuencia de la diferencia en incentivos es que, en sistemas basados en costos, la calidad a menudo es excesiva (otra ineficiencia asociada a Averch-Johnson) y que, en los sistemas basados en incentivos, la calidad se deteriora a los mínimos admisibles por el regulador.

Entre estos dos extremos, naturalmente aparecen los mecanismos regulatorios de poder intermedio, que combinan elementos de incentivos con elementos basados en costos³. Varios esquemas que clasifican como intermedios consideran fija/dada la infraestructura existente, la cual se debe remunerar. Luego, se estiman los costos que tendría una empresa que, a partir del instante de su tarificación, y por todo el período tarifario, se maneja en forma eficiente⁴. En principio, un mecanismo de poder intermedio podría ser más eficiente que los mecanismos puros, al permitir una mezcla óptima entre incentivos a reducir costos y aumentos excesivos de riesgo que elevan el costo de capital y, finalmente, las tarifas.

En el mundo real, de cierta forma se podría argumentar que la gran mayoría de los mecanismos son en efecto de poder intermedio, incluso en el caso de sistemas que, en teoría, son de alto poder. Las causas son dos: primero, el regulador corrige sus modelos si se apartan demasiado de los valores observados, ante el riesgo de haberse equivocado o por temor ante la posibilidad de hacer quebrar a la empresa; y, segundo, porque usualmente la empresa puede influir sobre los resultados regulatorios. El problema es que esa combinación usualmente no es producto de una comprensión previa del fenómeno, sino que se consigue mediante prueba y error en la implementación, como resultado de conflictos entre las empresas reguladas y el regulador.

2.2. Criterios de asignación de riesgo⁵

La teoría del riesgo señala que los criterios para asignar el riesgo en forma eficiente tienen dos componentes. Primero, existen los riesgos que son controlables, al menos parcialmente, por una de las partes del contrato. En tal caso, el riesgo controlable debe ser asignado al agente que puede controlarlo⁶. Por

³Entre ellos están los esquemas basados en menús de contratos de [13].

⁴Esta definición incluye a esquemas de tipo empresa modelo brownfield donde se consideran fijos los activos existentes, y se optimiza solo la expansión de la red para el período tarifario. Ver, por ejemplo, [11].

⁵Cuando nos referimos a riesgo, se refiere a variabilidad/incertidumbre de una variable o a la variabilidad producto de la variabilidad de las componentes de un resultado. En general este riesgo incluye el riesgo de colas, pero no se focaliza en este, que requiere herramientas como Value at Risk (VaR) o Conditional Value at Risk (CVaR).

⁶Este punto admite sutilezas, porque los costos pueden incluir partes controlables y no controlables que no se pueden separar fácilmente. Por ejemplo, usualmente los costos de

ejemplo, aquella parte del riesgo de costo de instalación que depende del buen manejo del privado o el riesgo en los costos operacionales, deben ser asignados a la empresa privada.

Segundo, existen riesgos no controlables, es decir, aquellos que son exógenos y no se ven afectados por las acciones de las partes. Un ejemplo es el riesgo de costo de insumos transados en mercados internacionales. Aquellos riesgos que no se pueden manejar deben asignarse según la capacidad relativa para absorber riesgo. La forma de asignación debe ser tal que se minimiza el costo total, considerando que el retorno exigido debido al riesgo no controlable eleva el costo de la empresa y que, en el caso de la autoridad regulatoria (que representa a los usuarios de los servicios regulados), el riesgo de tarifas produce desutilidad. Cuando se aplica este criterio a la asignación del riesgo no controlable, se debe tener en consideración que una empresa regulada usualmente recibe la mayor parte de sus ingresos del sector regulado, por lo que tiene una capacidad limitada para enfrentar riesgos no controlables. Por el contrario, los usuarios usualmente dedican una fracción relativamente baja de sus ingresos a un servicio regulado, por lo que tienen mayor capacidad para absorber este riesgo, especialmente si se distribuye en el tiempo.

2.3. Remuneración de la transmisión

A continuación, se enumeran tres alternativas para remunerar la transmisión eléctrica. Estas tres alternativas se escogen por su relevancia para la discusión en Chile. Estas corresponden a la valorización en base a: i) Valor Libro, ii) Valor Nuevo de Reemplazo, y iii) licitaciones.

2.3.1. Esquemas basados en Valor Libro

En estos esquemas, el costo que se debe remunerar, traspasándolo a tarifa, es la suma de los costos eficientes de operación, mantenimiento y administración (COMA), la depreciación, y el retorno a la base de activos regulados ($r \times RAB$), según se ilustra en la ecuación (1) (ver [8]). La base de activos regulados se construye como el valor depreciado real (no nominal) de los activos existentes, más los nuevos activos que se agregan según un plan eficiente de

construcción son asignados a la empresa según el criterio de asignación de costos controlables. Sin embargo, en ocasiones existe una componente no controlable. Por ejemplo, cuando se construye un túnel carretero, los costos dependen del tipo de roca. Independientemente de los esfuerzos de la empresa, a menudo la calidad de la roca solo se conoce en los puntos de sondaje. En tal caso, usualmente se parametrizan los costos en función del tipo de roca efectivamente encontrada. Esto implica que parte del riesgo de construcción se asigna a la parte (el Estado, o la sociedad) que mejor puede soportar el costo. Ver, por ejemplo, la Sección 4.a de [10].

inversiones para el período (aunque el horizonte de planificación es más largo). La tasa de retorno a utilizar se determina en el proceso regulatorio. El retorno a estos activos depende de la tasa de descuento apropiada a la industria y de los supuestos de la relación deuda-capital en la industria.

$$Tarifa = r \times RAB + Depreciacion + COMA \quad (1)$$

2.4. Esquemas basados en VNR

En este caso, las componentes que determinan el costo a remunerar son la anualidad del valor de inversión (AVI o AVNR), más los costos asociados al COMA, según se indica en la ecuación (2). Sin embargo, en cada período tarifario, se considera la base de activos existentes, pero valorizada al valor que costaría reproducirlos en forma eficiente al momento en que se realiza el proceso tarifario ⁷. Este valor está sujeto a vaivenes producto de variaciones en el costo de materiales, de la mano de obra, y de la tecnología constructiva, entre otros. Una vez obtenido este valor, más el de las inversiones adicionales requeridas, éstas reciben una remuneración que permite reproducir los activos considerando su vida útil y la tasa de retorno. La determinación del nivel del retorno requerido es usualmente parte del proceso regulatorio.

$$Tarifa = AVNR + COMA \quad (2)$$

Una variante de este esquema es la determinación de una empresa modelo que podría eventualmente prescindir de algunos de los activos existentes y así reducir la base de activos a remunerar. Este último esquema (considerando prescindibilidad de algunos activos) se utilizó para la tarificación del segmento de la subtransmisión en Chile con anterioridad a la reforma legal de 2015.

2.5. Esquemas basados en licitaciones

Los esquemas basados en licitaciones tratan de determinar algún parámetro relevante de la tarifa mediante ofertas competitivas. Por ejemplo, las empresas podrían ofertar directamente la tarifa a recibir durante un número de años específicos (i.e., ofertar AVI + COMA), el valor de la inversión (VI), o el monto total del valor presente de los ingresos por tarifa. En Chile, por ejemplo, se pueden observar los dos primeros tipos en obras nuevas y de ampliación para

⁷Esta es la diferencia con el modelo usado en distribución, en que la empresa modelo parte de cero (greenfield). En el caso de los activos de transmisión, se revalorizan los activos existentes.

la transmisión eléctrica, respectivamente. La tercera opción también es usada en Chile para el caso de las obras públicas como carreteras [5].

2.6. Planificación de la transmisión y su relación con la remuneración: Modelos centralizados vs descentralizados

Una adecuada planificación de las futuras inversiones en transmisión es clave para viabilizar un sistema eléctrico adaptado económicamente, confiable y sustentable. Cuando estas decisiones están en manos de las empresas, es esencial entregarles señales de eficiencia mediante el mecanismo remunerativo. Es en este contexto, donde aparecen mecanismos de incentivos para poder asegurar ciertos niveles de eficiencia en la inversión. Estos mecanismos pueden ser sencillos tipo empresa modelo y/o VNR, o más sofisticados tipo RIIO [17].

En un mundo en que las decisiones de inversión las toma una autoridad que actúa como planificador central, los incentivos anteriormente mencionados no son requeridos⁸. La razón es que si la autoridad central decide las ampliaciones y expansiones de la red con un criterio de costo-beneficio que busca maximizar el bienestar social (o minimizar el costo), no hay espacio para que las empresas sean incentivadas a tomar buenas decisiones de inversión. En este caso, donde la eficiencia productiva está asegurada por la autoridad de planificación, la remuneración cumple solamente el rol de remunerar los activos de forma suficiente⁹, sin necesidad de contar con un mecanismo que envíe señales de inversión a las empresas. Bajo este tipo de regulación, usar un valor histórico para las inversiones, y aplicar la tasa de retorno del periodo reduce el retorno requerido, al evitar introducir riesgo en la valoración de los activos. Esto aplica, también, para el valor de los activos de inversiones que son el resultado de licitaciones.

3. Tarificación por VNR: Condiciones y riesgos

La tarificación por VNR consiste en valorizar el costo de reproducir los activos existentes en la red en forma eficiente al momento en que se realiza el proceso tarifario. Es decir, se reconocen los activos de la red a valor nuevo de

⁸Ya que es la autoridad, y no la compañía, la que debe asegurarse de tomar las decisiones más eficientes. Ver, por ejemplo, [18].

⁹Mecanismos tarifarios deben cumplir con dos criterios básicos: suficiencia de ingresos y eficiencia productiva. La suficiencia de ingresos asegura que los costos se pueden cubrir con los ingresos percibidos para una empresa que es eficiente. Las tarifas también deben enviar las señales para que las empresas sean eficientes en la prestación del servicio (eficiencia productiva). Ver, por ejemplo, [9] o [15].

reemplazo. A continuación, analizamos las condiciones para que un esquema tipo VNR opere adecuadamente. Luego, discutimos sus principales riesgos asociados dichas condiciones. Condiciones requeridas para que un esquema de tipo VNR opere adecuadamente Dado que una empresa eficiente debe ser capaz de cubrir exactamente sus costos, es posible demostrar que existen condiciones para que la tarificación por VNR opere correctamente. Estas son:

- La demanda no crece o crece a una tasa conocida, sin incertidumbre, por lo que las inversiones eficientes son perfectamente previsibles (tanto por la empresa como por el regulador) y no hay posibilidad de que una instalación sea considerada no necesaria y retirada antes del fin de su vida útil.
- Que no haya cambios en los precios de los elementos utilizados en la construcción de las instalaciones, para que los VNR no cambien en forma aleatoria, introduciendo riesgo indeseado.
- Que no haya cambios tecnológicos que hagan que ciertas instalaciones sean consideradas no necesarias antes de lo presupuestado.
- Que los parámetros de vida útil, tasa de interés/retorno utilizada en la valorización y los costos usados en la valorización coincidan con los valores efectivos.
- Que el regulador conozca todos los parámetros del modelo de valorización y los utilice sin oportunismo, y que pueda hacer exigibles los estándares de calidad establecidos en la normativa.

Si se dan estas condiciones en un contexto donde la empresa tiene el poder de decidir inversiones, la empresa agregará inversiones en forma eficiente, porque no enfrenta el riesgo de variación en el valor de sus instalaciones, ni enfrentará riesgos asociados al retiro temprano de algunas de ellas antes de cumplir su vida útil. Si decide invertir más de lo óptimo, o con una tecnología costosa, algunas instalaciones podrían no recibir una remuneración que recupere su costo. Como se asume que la tasa de interés representa el costo efectivo de recursos (incluyendo apalancamiento y riesgo de tasas), la empresa no tiene interés en sobreinvertir. Por el contrario, si subinvierte, la calidad del sistema se deteriora y las multas por fallas de servicio están diseñadas para disuadir este comportamiento¹⁰.

Si no se cumplen todas estas condiciones, aún es posible que el esquema de valorización por VNR entregue incentivos apropiados para la inversión de la

¹⁰Es el significado de la frase: “pueda hacer exigible los estándares”.

empresa de transmisión, pero determinar si eso ocurre no es evidente. Algunos efectos podrían ir en la dirección de sobreinvertir (e.g., si la tasa de retorno es muy alta) y otros en la dirección contraria (e.g., si el riesgo es muy alto, aumentando el costo de capital exigible por los inversionistas), por lo que, solo por accidente, resultaría una combinación de efectos que neutralicen cualquier sesgo generado por la ausencia de las condiciones a.-e.

Nótese que la razón esencial para que tenga sentido introducir un mecanismo de VNR es porque es la empresa la que decide sus inversiones en condiciones de monopolio regulado. Cuando las inversiones en cada periodo tarifario no son decididas por la empresa, sino que forman parte de una planificación de largo plazo que busca maximizar el bienestar social, introducir riesgo en el valor de las inversiones ya realizadas no cumple ningún rol eficaz. En tal caso, el riesgo de retorno no produce ningún beneficio en términos de incentivos a tomar mejores decisiones. Peor aún, asignarle riesgos de mercado al valor de las instalaciones puede generar riesgo también para la sociedad (porque el valor de las instalaciones podría subir y con ello las tarifas). En esas condiciones, el mecanismo es equivalente a crear una lotería de suma cero, donde el riesgo no cumple ninguna función útil¹¹.

A continuación, se analizan algunos de los riesgos más relevantes en la tarificación por VNR.

3.1. Riesgos asociados a la tarificación por VNR

Esta sección se focaliza en la parte de inversión de la remuneración y en la determinación del VNR, dejando en un segundo plano los problemas relacionados con la estimación del COMA, sobre el cual pueden existir otros riesgos.

Riesgo por reemplazos/retiros de activos a destiempo con respecto a su vida útil regulada

La fórmula de pagos que se utiliza para calcular las tarifas está diseñada para que la remuneración recolectada en el tiempo a partir de dichas tarifas cubra, en teoría, y, de forma exacta, los costos de inversión en infraestructura (VI), más los costos asociados a la operación, mantenimiento y administración (COMA). A continuación, esta sección se focalizará en los primeros, es decir, los costos de inversión (VI). Los costos de inversión, además de incluir los materiales y equipos necesarios para prestar un servicio determinado de red, incluyen el costo que demandan los dueños del capital. Esto último se incorpora

¹¹Otro sector en que se han cometido estos errores regulatorios es en el caso de los peajes sombra utilizados en algunas carreteras concesionadas en Europa, ver [6].

mediante una tasa de retorno dentro de la fórmula para anualizar los costos de inversión. La fórmula de pagos requiere, además, una vida útil del activo que indica el tiempo de vida del activo y , por lo tanto, el tiempo por el cual se debe remunerar su servicio. Así, al pagar la anualidad determinada por la fórmula (3), llamada anualidad del valor de la inversión (AVI), durante la vida útil (VU) del activo, se recobrarán los costos de inversión (VI), más la remuneración r al capital del inversionista, ajustada por riesgo, según se muestra en (3).

$$AVI = VI \frac{r(1+r)^{VU}}{(1+r)^{VU} - 1} \quad (3)$$

A continuación, la Figura 1 muestra la remuneración acumulada (en valor presente) a lo largo del tiempo producto de un pago anual de \$9.45. Este pago anual, se determinó con la ecuación (3) utilizando un $VI = \$100$, un $r = 7\%$, y un $VU = 20$ años. Nótese que el pago anual (AVI) de \$9.45, cubre exactamente (en valor presente) el $VI = \$100$ al final del año 20. De la figura se desprende fácilmente que si un activo se remunera con un pago anual de AVI por un periodo más corto o largo con respecto a su vida útil regulada (VU), entonces el activo se estará sub-remunerando o sobre-remunerando, respectivamente. A continuación, se ahonda en las razones por las cuales un activo puede ser reemplazado de forma temprana o tardía.

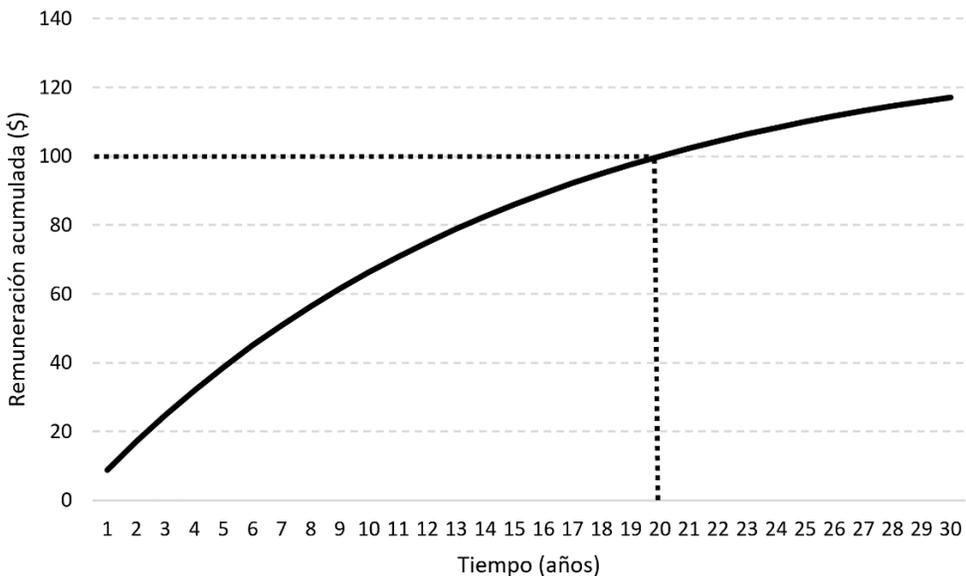


Figura 1: Remuneración acumulada en valor presente para un pago fijo anual.

Reemplazos tempranos

Como se mencionó, el pago AVI ha sido diseñado para remunerar de forma precisa el costo de inversión del activo, si este AVI se paga durante el número de años comprendido exactamente en la vida útil regulada. Existen, no obstante, situaciones que podrían llevar a un retiro o reemplazo temprano del activo. Es imperante mencionar que, en casos de retiros o reemplazos tempranos del activo, es decir, antes de terminar su vida útil regulada, la compañía incurre en una pérdida económica. Dependiendo de su naturaleza, estas pérdidas se pueden mitigar mediante seguros. A continuación, se discuten dos potenciales razones de retiros tempranos: por ampliaciones producto de la naturaleza creciente de la demanda por el servicio de transmisión, y por accidentes. Además, hay que tener presente que existe el riesgo de desconexión de líneas y subestaciones que no sean necesarias (o prescindibles) según lo que se establece en el Artículo 100 de la Ley General de Servicios Eléctricos (LGSE).

Con respecto a las razones por ampliaciones, existe el riesgo, por ejemplo, de que sea necesario ampliar una instalación específica por crecimiento de la demanda/generación y que parte de la instalación existente tenga que ser retirada. En ambos casos, existe un término anticipado de la vida útil que llevan a una sub-remuneración del activo existente. Por ejemplo, este es el caso del Tap Alto Melipilla 2×110 kV de propiedad de Chilquinta (instalado en el año 1996) que quedaría fuera de servicio debido a la instalación de la nueva línea Lo Aguirre – Alto Melipilla 2×220 kV.

Con respecto a las razones por accidentes, existe un sinnúmero de situaciones como, por ejemplo, la ocurrencia de desastres naturales, que afectan directamente la integridad de la infraestructura eléctrica [20, 16]. Terremotos, por ejemplo, tienen la capacidad de dañar activos en líneas y subestaciones. Incendios forestales causados por terceros pueden también cortar conductores y dañar torres (Figura 2). Aunque varias de estas situaciones pueden ser cubiertas por seguros, existen otras que no. Por ejemplo, según información proporcionada por empresas de redes, no existen en Chile seguros que cubran catástrofes que afecten líneas de transmisión (solamente subestaciones)¹². Estas situaciones, por lo tanto, pueden producir un daño económico importante en compañías de redes, cuando no se les reconoce dicho reemplazo anticipado.

Con respecto a las razones socio-ambientales, existe una creciente conciencia de las autoridades y de las comunidades locales acerca de las interacciones entre los activos de transmisión y el medioambiente [14]. Esto no es un fenómeno nacional, sino que ha ido ocurriendo paulatinamente en varios

¹²Además, las Bases del estudio tarifario, en su Capítulo II, no permite la incorporación de seguros de líneas de transmisión en la función de costos de la empresa eficiente.



Figura 2: Daño en línea y subestación del SEN por terremoto, 27 Feb 2010.

países. Aunque esto generalmente es un problema respecto a nuevas instalaciones, también ha comenzado a afectar instalaciones existentes. Por ejemplo, razones socio-ambientales pueden llevar a soterrar una línea o a trasladarla, como es el caso del Bypass para la línea 1×220 kV Atacama – Esmeralda, línea 1×110 kV Esmeralda – La Portada y Línea 1×110 kV Mejillones – Antofagasta. Este último caso se originó por la ubicación de campamentos bajo la línea. A futuro, podrían aparecer presiones para disminuir la contaminación auditiva (ruido) de algunas líneas, lo cual se puede realizar mediante cambios en los conductores (y, probablemente, torres). También, podrían aparecer presiones para mejorar la apariencia visual de algunas instalaciones, llevando a un recambio de torres.

En estos casos, si bien los nuevos activos podrían ser adecuadamente remunerados, el reemplazo o retiro anticipado de los activos antiguos podría producir una pérdida económica para la empresa en el caso de no aplicarse compensaciones.

Reemplazos tardíos

Como se mencionó, el pago del AVI ha sido diseñado para remunerar de forma precisa el costo de inversión del activo si se paga por los años correspondientes a la vida útil regulada (ver Figura 1). Existen, no obstante, situaciones que podrían llevar a un retiro o reemplazo tardío del activo. Es imperante mencionar que, en casos de retiros o reemplazos tardíos del activo, es decir, después de terminar su vida útil regulada, la compañía podría incurrir en una ganancia económica extraordinaria.

Dicho reemplazo tardío se da principalmente porque el activo puede seguir

sirviendo y prestando el servicio para el cual fue diseñado más allá de su vida útil regulatoria. El estado de salud y la degradación del equipo corresponden a procesos estocásticos y la vida útil no corresponde un valor totalmente predecible. Además, la vida útil depende del uso que se le da al activo. Así, un activo (transformador, por ejemplo) que ha sido utilizado a altas cargas de potencia (y por lo tanto a altas temperaturas), tendrá una vida útil menor que el mismo activo siendo utilizado a una potencia más baja. Una buena política de inspecciones y mantenimientos puede retrasar el proceso de degradación del equipo, aumentando su vida útil (ver, por ejemplo,[1, 2]).

El actual modelo regulatorio, con un AVI que se paga de forma constante e independiente de cuando se realice el reemplazo, crea incentivos para que dicho reemplazo sea óptimo, utilizando de forma eficiente la vida útil efectiva de los activos. Al desacoplar los ingresos de los costos, la empresa está incentivada a minimizar sus costos, en este caso, compuestos por los costos de inversión asociados al reemplazo y los costos de multas asociados a deficiencias en la calidad de suministro. Si las multas reflejan *correctamente* el costo de falla que sufre la sociedad, las decisiones de reemplazo de la empresa serán las mismas decisiones que las que adoptaría un planificador social benevolente. Por ejemplo, si un activo tiene una tasa de falla que se traduce en multas con un valor presente esperado inferior al costo de reemplazo, la empresa lo mantendrá con el fin de minimizar costos¹³. Si las fallas del mismo activo aumentan de modo que el valor presente esperado de las multas es superior al costo de reemplazo, la empresa debería reemplazarlo para minimizar sus costos totales.

Si bien los incentivos del actual modelo son los correctos desde una perspectiva del momento del reemplazo, en la práctica se sobre-remunera a la empresa. Esto se debe a que la vida útil o tasa de reemplazo asumida en la regulación es inferior a la vida útil efectiva. Incluso si los montos de sobre-remuneración por este criterio son pequeños (como usualmente lo argumentan las empresas reguladas), su existencia da pábulo a que el regulador intente ajustar la rentabilidad de los activos en general, distorsionando aún más el esquema regulatorio. La Sección 4.3 discute como tratar este problema.

3.1.1. Riesgo por fluctuaciones de precio

Según el artículo 71-10 de la Ley Eléctrica, el VI de las instalaciones de transmisión se compone de los costos de adquisición e instalación de sus componentes, valorados a precios de mercado, con la excepción de los derechos

¹³Este costo de reemplazo debe incluir el valor descontado de las multas asociadas al activo nuevo.

asociados al uso de suelo y el costo de servidumbres, que se consideran a sus valores originales, indexados según IPC. Esta definición del VI genera varios riesgos para las empresas (y, por contrapartida, para los consumidores), los que se considerarán en lo que sigue.

Para esto dividimos los costos en aquellos que favorecen a la empresa porque se elevan en el tiempo; aquellos costos cuyo efecto es ambiguo, porque no hay una tendencia clara; y aquellos que favorecen al regulador (o a los usuarios), pues reducen los costos de equipos y de instalación. En el grupo de los que se elevan en el tiempo, se encuentran los costos laborales unitarios y los costos ambientales y similares. En el grupo cuyos efectos son ambiguos, se encuentra los costos de materiales, especialmente cobre y aluminio. Finalmente, en el grupo que favorece la caída en los VNR están los cambios tecnológicos, que pueden incluir aquellos que ahorran costos laborales o aquellos que potencialmente pueden cambiar la tecnología utilizada en la industria de la transmisión.

Riesgo de costos de terrenos y servidumbres

Los riesgos de costos de terrenos y servidumbres se pueden excluir del análisis. Este riesgo no existe, dado que la Ley no reconoce los cambios en los valores reales de estas componentes del VI¹⁴. Sin embargo, este mismo hecho confirma que usar el mecanismo de VNR tiene poco sentido: ¿por qué usar los valores de mercado de materiales, pero mantener los valores originales de terrenos y servidumbres? ¿No sería mejor mantener todo a los valores de compra, dado que no tiene efectos sobre las acciones de los actores?.

Riesgo de costo por efectos socioambientales

Estos costos son crecientes, porque es cada vez más difícil instalar equipos de transmisión dada la oposición de grupos ambientales y otros que se oponen a los proyectos de transmisión. Estos costos incluyen compensaciones a comunidades, posibles medidas para reducir el impacto visual, traslados, etc. Dado

¹⁴ “El V.I. de una instalación de transmisión es la suma de los costos eficientes de adquisición e instalación de sus componentes, de acuerdo con valores de mercado, determinado conforme a los incisos siguientes. En el caso de las instalaciones existentes, el V.I. se determinará en función de sus características físicas y técnicas, valoradas a los precios de mercado vigentes de acuerdo a un principio de adquisición eficiente. Sin perjuicio de lo anterior, respecto de los derechos relacionados con el uso de suelo, los gastos y las indemnizaciones pagadas para el establecimiento de las servidumbres utilizadas, para efectos de incluirlos en el V.I. respectivo se considerará el valor efectivamente pagado, indexado de acuerdo a la variación que experimente el Índice de Precios al Consumidor.” (Art 103°, incisos tercero a quinto, LGSE).

que son costos de instalación (y no de servidumbres ni de terrenos), deberían estar considerados en el VI de los equipos, y en promedio deberían subir con el tiempo. Es posible que la interpretación del cuarto inciso del artículo 103° se extienda a estos costos, pero no es una interpretación evidente cuando se refiere a compensaciones no asociadas a servidumbres, por ejemplo, modificaciones ineficientes de trazado para responder a oposición local, o cuando se trata de medidas para reducir el impacto visual, auditivo y similares. Otro tipo de costos son los retrasos en los proyectos. Por ejemplo, la construcción de la reciente línea Cardones-Polpaico 2×500 kV, tuvo retrasos de más de 17 meses. Esto resultó en pérdidas cuantiosas para los participantes del mercado (además del propietario de la línea).

Costos de mano de obra

Otro de los costos de instalación es el costo de la mano de obra, el cual se ha ido elevando en el tiempo, como lo muestra la Figura 3. Esta componente es importante, porque incide a través de todos los costos, incluyendo los de instalación de faenas. En este caso, el efecto es claro: favorece a la empresa porque su valor aumenta de forma sistemática.

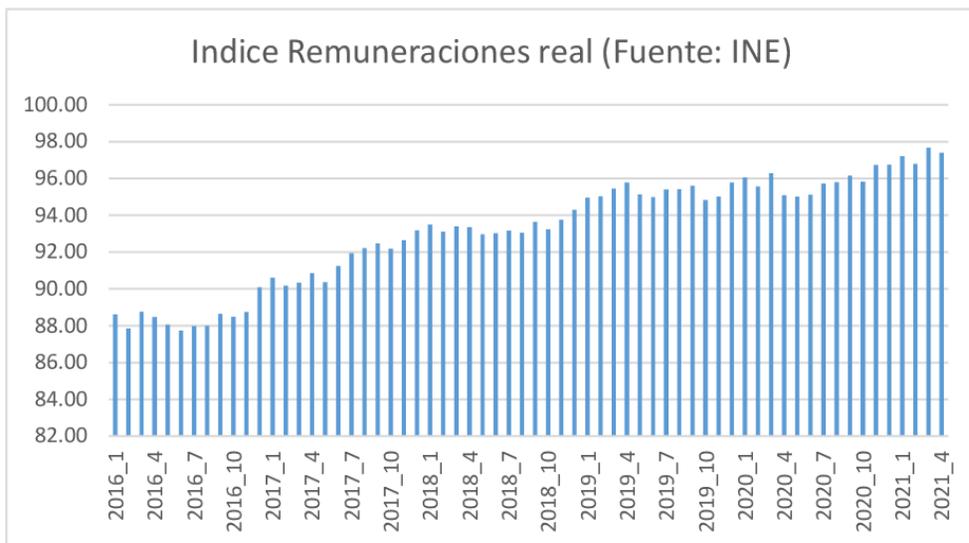


Figura 3: Índice de remuneraciones en Chile. Fuente: INE.

Costos de los materiales

Los costos de los materiales son variables, como es bien sabido. La Figura

4 muestra el precio diario del cobre en COMEX de los últimos 10 años.



Figura 4: Precio diario del cobre Comex (USD/lb). Fuente: Macrotrends.

La variación en los precios del aluminio, sustituto del cobre, es similar y tiene bastante correlación con el precio del cobre, como puede observarse en la Figura 5. El costo de estos metales es una de las principales componentes de los costos de los equipos y, como se observa, puede favorecer al regulador o a la empresa; no obstante, dependiendo del año, puede tener un efecto importante, positivo o negativo, en el valor de los materiales. Por lo demás, la forma en que se transan los materiales en la industria, aunque depende del costo de los metales, no es directa: el costo de los conductores no necesariamente es igual al costo de materiales más un costo fijo de procesamiento, sino el resultado de una negociación entre las empresas proveedora y compradora.

Efectos tecnológicos

Existen dos tipos de efectos que podrían afectar los costos de instalación y de materiales. Primero, están aquellos cambios tecnológicos que ahorran trabajo y que, por lo tanto, pueden neutralizar o incluso revertir el efecto de aumento de costos de mano de obra. Este efecto tiende a favorecer al regulador.

El segundo tipo de cambio tecnológico es aquél que altera totalmente la industria, como ocurriría si se descubren mejores superconductores de alta temperatura que se pueden usar en la industria de transmisión a menor cos-

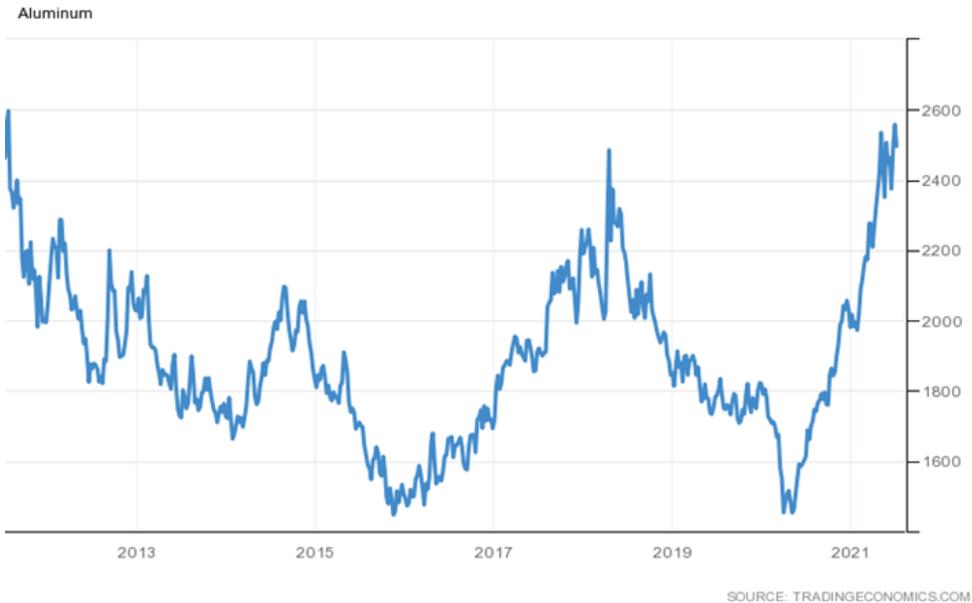


Figura 5: Precio diario del aluminio (USD/ton), LME. Fuente: Macrotrends.

to. Esto produciría una menor demanda, y un menor precio, de tecnologías antiguas. En ese caso, el efecto bajo un esquema VNR es que algunos de, o eventualmente todos, los activos ya instalados podrían presentar un valor sustancialmente menor, percibiendo una remuneración menor.

3.1.2. Riesgos/desafíos en la determinación de los precios de mercado

Uno de los problemas que enfrenta el regulador es el de determinar los precios de los distintos elementos necesarios para calcular los costos de instalación y de materiales. No hay un mercado tradicional para muchos de los componentes (por ejemplo, conductores o transformadores), sino que las empresas solicitan cotizaciones y negocian con los proveedores un precio final. El poder de negociación de las empresas puede ser distinto (porque las cantidades requeridas varían), y puede cambiar en el tiempo. Incluso si la empresa regulada entregase esa información cuando la tuviese, esta correspondería a una transacción pasada, es decir, a un valor histórico, y probablemente no corresponde a la situación del momento regulatorio. Pueden existir diferencias, también, en la evaluación de los costos de instalación de faenas, costo y número de trabajadores, intereses intercalarios, etc. No hay formas claras de resolver este problema si se intenta determinar los precios a partir de las componentes de costo, como lo indica la Ley. Bajo el sistema chileno, se con-

trata un estudio para realizar la valorización usando bases técnicas (que se discuten con los participantes) para realizar este proceso. Pero luego las partes contratan especialistas, expertos que conocen los precios de los distintos elementos, cuyos resultados se usan para presentar observaciones a los valores obtenidos en el estudio, y, a menudo, aparecen diferencias importantes en los valores considerados. Dados estos problemas, el resultado de la valoración del VNR no es solo producto del estudio, sino que los resultados se mueven según la capacidad de presión de los distintos actores. Si bien, el Panel dictamina en casos de discrepancias importantes, éste no dispone de la información clave necesaria para determinar con certeza los valores cuestionados.

En la Sección 4.2 ilustramos una propuesta que permitiría resolver esta problemática, eliminando la dependencia del estudio de mercado de precios, el cual presenta niveles de riesgo muy elevados para las partes (inversionistas y consumidores).

3.1.3. Riesgos asociados a un reconocimiento inadecuado de la tasa de retorno

Con el fin de determinar la tasa de retorno de compañías monopólicas reguladas de transmisión, varios países del mundo utilizan el modelo CAPM¹⁵. Este modelo permite ajustar las ganancias de las empresas reguladas según las condiciones reales de mercado. Esta fórmula (descrita a continuación) establece que la tasa de retorno sobre los activos de la empresa regulada debe ser igual a la tasa libre de riesgo, más el producto del denominado riesgo sistemático, multiplicado por el premio por riesgo de mercado.

$$r = r_f + \beta \times PRM^{16} \quad (4)$$

Como se puede observar, esta fórmula contiene solamente tres parámetros a estimar¹⁷. Los fundamentos técnicos dictan que la estimación de cada parámetro debiera seguir los siguientes principios:

¹⁵Australia, Brasil, Colombia, México, Reino Unido, España, Holanda y los países de la Unión Europea, entre otros.

¹⁶Esta fórmula para la tasa de retorno incorpora términos como el riesgo país, y otros factores específicos de riesgo que se agregan de manera aditiva al final de la fórmula. Por simplicidad, estos factores se omiten de nuestro análisis debido a que son independientes del mayor riesgo específico a la industria derivado del uso de regulación por VNR. Los inversionistas de la transmisión eléctrica en Chile enfrentan un riesgo mayor que el de países comparables debido a la regulación por VNR.

¹⁷Por simplicidad, esta explicación se basa solamente en los tres parámetros descritos, pese a que cada uno de estos tres parámetros puede depender de otros (e.g., tasa de mercado, tasa de impuestos, etc.).

1. Para r_f : La tasa libre de riesgo, que refleja la tasa de rentabilidad (esperada a futuro) para un activo libre de riesgo en Chile durante el periodo tarifario (e.g. próximos 4 años, si es que el cálculo se realiza al comienzo del periodo tarifario).
2. Para β (beta): El riesgo sistemático, que refleja el perfil de riesgo de la empresa regulada durante el periodo tarifario y es cuantificado en función de la dependencia entre la rentabilidad de un activo de la empresa regulada y la denominada tasa de mercado, que, en este caso, refleja una tasa de rentabilidad “promedio” para el mercado chileno. A mayor dependencia, mayor beta.
3. Para PRM : El premio por riesgo de mercado, que refleja la diferencia entre la tasa de mercado y la tasa libre de riesgo durante el periodo tarifario.

Uno de los principales desafíos en este ejercicio (y, además, una de las principales desventajas del modelo CAPM) es la determinación de estos tres parámetros, sobre los cuales hay un nivel de incertidumbre. Para los numerales anteriores 1 y 3, existe, por ejemplo, una dependencia importante con respecto al conjunto de instrumentos financieros seleccionados para hacer el cálculo. Esto quiere decir que el resultado es dependiente de una serie de decisiones arbitrarias como, por ejemplo, si se selecciona un bono BCU o BTU ¹⁸ a 10 o 20 años como representativo de la tasa de libre de riesgo. Tampoco es claro el periodo sobre el cual se debe calcular el promedio histórico de estas rentabilidades (¿últimos seis meses? ¿último año? ¿últimos 5 años?) y cómo se pueden eliminar o “filtrar” desviaciones producto de eventos coyunturales que no son representativos de la conducta general del activo.

El riesgo sistemático (o beta), que refleja el perfil de riesgo que enfrenta el inversionista, es determinado mediante un estudio comparativo con otras compañías reguladas (cuyo beta es conocido) de similar perfil de riesgo. En el caso chileno, este parámetro es uno de los que ofrece mayores desafíos en su cálculo, pues, ¿qué otro sector (cuyo beta sea conocido) está expuesto a los mismos niveles de riesgo que el sector de transmisión eléctrica en Chile, el cual es regulado mediante una metodología específica como la del VNR? En la metodología de cálculo, se utiliza típicamente un conjunto de compañías reguladas del mismo sector o de sectores similares. No obstante, como se mencionó anteriormente (y se profundizará a continuación), el riesgo de negocios regulados en el mundo (desarrollado) presenta niveles de riesgo menores debido a

¹⁸BCU son bonos emitidos por el Banco Central de Chile mientras que los BTU son bonos de la Tesorería General de la República de Chile.

marcos regulatorios acorde con el del negocio que regulan.

Más aún, mientras gran parte de las empresas de redes eléctricas en el mundo son reguladas mediante principios que permiten limitar los niveles de riesgo a los cuales se encuentran expuestas, por ejemplo, utilizando metodologías de remuneración que reflejen (al menos en parte) la base real de activos de la empresa (RAB o Regulated Asset Base por su sigla en inglés) y su valor de libro¹⁹, en Chile el principio es ignorar los costos de la empresa real y exponerla a un perfil de riesgo más competitivo, haciéndola (de hecho) competir contra los actuales precios de mercado, los cuales son finalmente los que se reflejan en las tarifas. Esto produce que una empresa de transmisión en Chile quede expuesta a variaciones importantes en sus ingresos, que no reflejan realmente los costos históricos del sector.

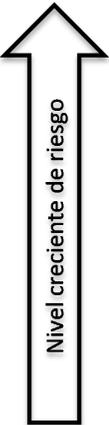
Esto es diametralmente opuesto a lo observado en los sectores de redes eléctricas en países como Australia, Estados Unidos, Finlandia, Francia, Grecia, Polonia, Rumanía, Eslovaquia, República Checa, Colombia, Brasil y, especialmente, en el Reino Unido (solo por nombrar algunos), el cual ha llevado a cabo recientemente mejoras significativas en la regulación de las redes eléctricas. La regulación del Reino Unido, llamada RIIO, es interesante, además, porque reconoce en el cálculo de la tasa de retorno los riesgos asociados al marco regulatorio y a las fórmulas particulares de pago a los cuales se acogen las distintas empresas. Por ejemplo, una empresa puede decidir, ex ante, pasar una mayor parte de sus costos futuros a tarifa, tomando menos riesgos. En estos casos, la empresa recibe un menor premio por riesgo. En el caso contrario, si la empresa decide tomar más riesgos, con una tarifa menos responsiva a los costos reales de la empresa a futuro, ésta recibirá un aumento en su tasa de retorno. Esto demuestra como regulaciones avanzadas han reconocido de forma explícita la relación que existe entre el marco regulatorio, sus riesgos para los inversionistas, y la tasa de retorno regulada. Más aún, es reconocido que la regulación por Valor Nuevo de Reemplazo o VNR crea niveles de riesgos para los inversionistas que van más allá del común de los negocios regulados. En algunos textos regulatorios, de hecho, el riesgo asociado a este tipo de regulación es catalogado como inaceptable para los inversionistas. Todo lo anterior significa que no es correcto utilizar el beta de otras empresas reguladas para aplicarlo a una empresa eléctrica en Chile, inhabilitando en la práctica usar la metodología CAPM. El perfil de riesgo que enfrenta una empresa de transmisión en Chile sería, por diseño, el de una empresa que enfrenta competencia, ya que el mismo diseño regulatorio así lo establece. A continuación, a modo de referencia, se presentan los betas y las tasas resultantes de retorno ilustrativas

¹⁹Por ejemplo, la estructura tarifaria predominante en Europa es en base a RAB. Ver [19].

de varias actividades reguladas y no reguladas en varios países del mundo (incluyendo Estados Unidos, Canadá, México, Brasil, y países de Europa y Asia). A modo ilustrativo, la Tabla 1 muestra que las variaciones del parámetro beta, que influye directamente en el cálculo de la tasa de retorno, puede presentar variaciones de hasta más de 4 veces, desde valores como 0.32 hasta 1.42, dependiendo del nivel de riesgo de la actividad. Así, bajo una regulación como el VNR, es probable que sea más apropiado utilizar el beta asociado a una industria competitiva que el valor asociado a una industria que es regulada como un monopolio más convencional.

Tabla 1: Niveles de riesgo sistemático (o beta) y tasa de retorno ilustrativa (calculada como $3\% + \beta \times 7\%$)²¹.

Industria	Tamaño de la muestra	Beta (desapalancado)	Tasa de retorno ilustrativa según beta
Industria química	6	1.42	12.97%
Computación (hardware)	57	1.41	12.89%
Industria farmacéutica	237	1.32	12.26%
Retail (Online)	79	1.30	12.09%
Software (Internet)	44	1.25	11.78%
Tabaco	17	1.09	10.60%
Metales y minería	94	1.01	10.05%
Material de construcción	42	0.88	9.15%
Energía convencional (carbón)	23	0.78	8.43%
Energías verdes	21	0.77	8.41%
Restaurant	78	0.63	7.41%
Transporte aéreo	18	0.61	7.27%
Distribución de petróleo y gas	20	0.61	7.26%
Hospitales	34	0.55	6.85%
Utility (agua)	19	0.32	5.25%



En conclusión, la metodología de CAPM tiende a subestimar de forma sistemática el valor efectivo en la tasa de retorno de una empresa de transmisión en Chile, debido a que no se considera la diferencia entre el riesgo en una industria cuyas tarifas son reguladas por VNR y una regulada con los métodos usuales en la industria de transmisión. Esto se debe a que no se reconoce el nivel de riesgo efectivo que enfrentan los inversionistas de redes eléctricas en Chile producto de las particularidades de la regulación por VNR. Es importante mencionar que, aunque la fórmula del CAPM va acompañada de un valor piso y techo (7 y 10% respectivamente después de impuesto), esto no elimina su sesgo sistemático en la estimación de la tasa de retorno a remunerar. Como

²¹Estas tasas ($r_f = 3\%$ y $PRM = 7\%$) debieran variar por país, pero, para efectos ilustrativos, se ha escogido valores constantes. Datos de betas disponible en [4].

se ha mencionado anteriormente (y se ahonda en la Sección 4), el riesgo adicional que enfrentan las empresas reguladas debido a la regulación por VNR dejó de ser justificable al comenzar la planificación central de la transmisión. El hecho de seguir utilizándolo es un error regulatorio, que se ve amplificado cuando se utilizan betas de países en que estos riesgos no existen.

3.1.4. Riesgos regulatorios

Los procesos de regulación de tarifas están influenciados por presiones de distintos grupos de interés. Las empresas desean que los valores sean los más altos posibles, y, el regulador, como agente de los usuarios, desea que sean lo más bajas posibles. En general, un modelo regulatorio intenta que el equilibrio de este juego se acerque lo más posible al mínimo valor que permite que la empresa realice las inversiones necesarias para ofrecer la calidad de servicio definida en la normativa. El resultado de estas presiones es un valor con cierto riesgo, pero es un tipo de riesgo conocido por las empresas reguladas y por el regulador (así como por los análisis de la literatura especializada). En el caso de procesos regulatorios bien concebidos, este riesgo no debería representar un costo excesivo para las partes. Los costos reales aparecen cuando una de las partes aprovecha alguna condición o deficiencia del sistema para obtener una renta sobrenormal, en el caso de la empresa o una apropiación injusta, en el caso del regulador.

Efectos que favorecen la aparición de rentas

A menudo, cuando ciertos aspectos regulatorios no están bien concebidos, se prestan para la aparición de rentas indeseadas. Reiterando un ejemplo anterior, nuestro procedimiento de remuneración de instalaciones y equipos mantiene la remuneración de una instalación pese a haber cumplido su vida útil, pese a que ya se han pagado las anualidades que costean el costo de inversión. El objetivo es que una instalación que aún presta servicios en forma eficiente siga utilizándose y no se le reemplace en forma adelantada. Esto ocurre porque, a menudo, las vidas útiles contables no son equivalentes a las vidas útiles económicas (como se describen en las secciones 3.2.1.2 y 4.3 sobre el tema).

Un esquema que continúa pagando la anualidad cuando la vida útil se ha cumplido genera una sobre-renta, porque el valor de los activos ya fue pagado. En general, las vidas útiles económicas solo las puede determinar la empresa, porque un mismo equipo con la misma antigüedad puede estar mucho menos solicitado en una zona que en otra, o estar en mejores condiciones o

simplemente porque la vida útil promedio de ese activo está subestimada por la autoridad regulatoria. Por lo tanto, pese a ser equivalentes en características nominales y antigüedad, tienen diferentes vidas útiles efectivas. A la empresa se le debe ofrecer un ingreso por conservar el activo hasta que sea económicamente eficiente su reemplazo, pero esta renta puede ser menor que el AVI (esto se discute en la Sección 4.3).

Como se ha mencionado antes, uno de los problemas de la aparición de sobre-rentas es que tienden a deslegitimar el procedimiento regulatorio. Esto puede llegar a producir una respuesta del regulador en la forma de una expropiación regulatoria, que se describe a continuación.

Expropiación regulatoria

La expropiación regulatoria consiste en la apropiación de parte del retorno esperado del concesionario luego que este ha realizado sus inversiones y no tiene la opción de retirarlas. Esto puede ocurrir por presión política a favor de los usuarios, cuando el regulador no posee suficiente independencia. Un ejemplo es lo que ocurrió con las tarifas de servicios públicos en Chile hasta antes de la implementación de la regulación por incentivos (es decir antes de la década de 1980), que usualmente solo pagaban los costos operacionales, impidiendo que las empresas de telecomunicaciones, por ejemplo, invirtieran y prestaran un servicio adecuado a la demanda²².

Algunos elementos de expropiación regulatoria recientes son los préstamos forzosos que las empresas de distribución han debido incurrir por las leyes que prohíben el corte de servicio sin compensación por intereses por mora. Esas medidas aumentan el riesgo percibido de las inversiones y, por lo tanto, el retorno que se requiere para invertir.

Con el fin de evitar confusiones, se debe señalar que una multa por mala calidad de servicio (violando la norma técnica) no corresponde a una expropiación regulatoria. En cambio, si hay expropiación regulatoria si la normativa de calidad se ha hecho más exigente, pero no se ha elevado la remuneración a los activos para permitir las inversiones que requiere la nueva calidad de servicio.

²²En [7] se señala que con anterioridad a 1973: “*El Estado y el público se quejaban continuamente de la obsoleta tecnología de la CTC, debido a la lentitud para establecer nuevas tecnologías de larga distancia. Por su parte, la CTC se quejaba porque su rentabilidad era menor a la estipulada en el contrato-ley*”, p.8.

4. Discusión y propuestas

La tarificación por VNR consiste en utilizar precios de mercado en lugar de los precios/costos efectivamente incurridos (e.g., Valor Libro) para valorizar las instalaciones. Específicamente, la valorización de una instalación consiste en la suma de los productos entre el vector de precios y el vector de cantidades de las distintas componentes de la instalación. El mecanismo de VNR utiliza un vector de precios distinto de los costos efectivamente incurridos, ya que usa los precios de mercado en el momento del control regulatorio, y asume activos nuevos. Aunque este marco crea riesgos para las empresas reguladas (i.e., incurrir en costos mayores a los reconocidos en tarifa), la ventaja de este marco tarifario es que incentiva a la empresa a tomar decisiones de expansión de forma eficiente, comprando al mejor precio posible y seleccionando las tecnologías más eficientes para entregar el servicio de transporte deseado. No obstante, en Chile, las modificaciones de la regulación de la transmisión han ido entregando las decisiones de inversión de los segmentos nacional y zonal a un planificador central, la Comisión Nacional de Energía. En este contexto, es relevante reflexionar acerca del rol que cumple realmente el VNR en la eficiencia del sector.

En el marco en donde las decisiones de expansión permanecen con la empresa, la valorización por VNR traspasa riesgos a la empresa, lo que incentiva la inversión eficiente. La empresa debe invertir cuidadosamente porque los activos serán valorizados en el futuro a precios competitivos de mercado, distintos a los costos efectivamente incurridos. Como se explicó en la Sección 2, en este caso, el traspaso de riesgo a la compañía cumple un rol esencial: promover eficiencia en la inversión para la provisión del servicio de transporte eléctrico y simular el resultado de una industria en competencia. En el actual marco, en cambio, las decisiones de inversión son centralmente planificadas. Por lo tanto, el traspaso de riesgo a las empresas solo crea un alza en el costo del capital, lo que finalmente se traspasa a precio. Por lo tanto, es posible concluir que la aplicación de la tarificación por VNR no crea beneficios cuando se acopla a planificación central y, más bien, crea ineficiencias. El consenso internacional es que, si las decisiones de inversión son centralmente planificadas, se debería limitar el traspaso de riesgo en la valorización de sus activos. De esta forma, se pueden alcanzar menores costos de capital, manteniendo la eficiencia tanto en el diseño de la red (planificación centralizada) como en la construcción de las instalaciones (dado que se trata de licitaciones abiertas).

Un argumento a favor de la regulación por VNR actual se relaciona con los reemplazos. Estas decisiones se mantienen en manos de las empresas, por lo que el VNR entrega incentivos a un reemplazo eficiente. Dado que desacopla los ingresos de los costos, la empresa está incentivada a minimizar sus costos. Hay dos costos que la empresa debe contrastar: los costos de inversión asociados al reemplazo, y los mayores costos por multas asociados a las fallas en la calidad de suministro sumados a los mayores costos de mantenimiento y operacionales de los equipos. Si las multas reflejan adecuadamente el costo de falla que sufre la sociedad, las decisiones de reemplazo de inversión de la empresa serán idénticas a las decisiones de un planificador social benevolente plenamente informado.

Este no es el mejor mecanismo para conseguir el objetivo de un reemplazo eficiente. Se pueden replicar sus niveles de eficiencia en las decisiones de reemplazo a un costo menor para los usuarios. En la siguiente sección, se propone un mecanismo sencillo que corrige el problema de los reemplazos sin depender de una tarificación por VNR que altera los ingresos en forma aleatoria cada 4 años. Además, el beneficio por reemplazo eficiente no compensa los mayores costos debido al riesgo de valorización de activos que genera el mecanismo de VNR.

A la luz de los antecedentes anteriores, existe suficiente evidencia para recomendar los siguientes cambios en el esquema actual de VNR:

- Minimizar el traspaso de riesgos a la empresa producto de una valorización aleatoria del VI cada 4 años en un modelo con decisiones de inversión centralizadas.
- Minimizar la sobre-renta asociada a instalaciones que se siguen pagando más allá de la vida útil definida por el regulador.

Las dos recomendaciones anteriores llevan a una tercera recomendación asociada al valor inicial (VI) de los activos existentes. A continuación, se desarrollan estas tres recomendaciones con mayor detalle.

4.1. Minimizar los niveles de riesgo a traspasar a la empresa

Con los antecedentes anteriores, proponemos modificar la Ley para reducir la incertidumbre del actual sistema de valoración de activos. La razón de esta recomendación es que, como se ha señalado anteriormente, cuando existe planificación central de la transmisión, ese riesgo no mejora los incentivos a realizar inversiones eficientes (más allá de la minimización de costos de las inversiones requeridas por el planificador), pero aumenta los costos de capital. Si la valoración de activos ya no depende de su valor nuevo de reemplazo y solo

enfrenta riesgos debido a la variación de tasas de descuento en el periodo, el retorno requerido de las inversiones es menor. La primera propuesta consiste en fijar un valor inicial para el VI de las instalaciones existentes, manteniendo dicho valor en el tiempo.

Para definir los valores iniciales de VI, proponemos utilizar como referencia los resultados de las licitaciones de las obras de ampliación, como se explicará en la siguiente subsección.

Con respecto a la duración de los pagos, las instalaciones solo percibirían la anualidad hasta el completar su vida útil regulatoria. Por lo tanto, se debe establecer la antigüedad de las instalaciones existentes en el VI del año base, lo cual en otros países ha requerido una negociación compleja²³. El nuevo mecanismo de valoración tiene menor riesgo, lo que permitiría que las empresas aumenten su apalancamiento y aumenten su retorno, un factor que se debe considerar al estimar las tasas de retorno.

Tanto la determinación de los VI iniciales, como el mecanismo para minimizar las sobre-rentas post vida útil, se detallan a continuación.

4.2. Estimación de los valores de inversión iniciales: Uso de resultados de licitaciones

Existe información de dos tipos de licitaciones que se podría utilizar para determinar los precios de mercado y así valorizar los activos existentes: las licitaciones de obras nuevas y las de obras de ampliación. Alternativamente, en lugar de eliminar el estudio de mercado, los resultados de las licitaciones se podrían utilizar para desarrollar cotas entre las cuales podrían ubicarse los resultados del estudio, lo que daría más confiabilidad a sus resultados. A continuación, se desarrolla la idea de utilizar los resultados de las licitaciones para construir los VIs de los activos existentes. En el caso de obras nuevas, el valor obtenido en la licitación es el Valor Anual de la Transmisión (VATT). Este valor es la menor suma de la Anualidad del Valor de Inversión (AVI) más los costos de operación y mantenimiento (COMA) ofrecidos en la subasta. Este VATT se mantiene constante durante 5 períodos tarifarios; luego la obra pasa al régimen general de VNR.

En el caso de obras de ampliación, la empresa constructora ganadora es la que ofrece el menor Valor de Inversión (VI), el cual es remunerado por la empresa de transmisión. Este VI permanece constante por cinco periodos tarifarios en el cálculo de la tarifa (AVI+COMA). La anualidad del valor de

²³Una posibilidad sería que las empresas de transmisión pudieran entregar la edad de sus instalaciones en forma creíble (lo que se podría verificar mediante un muestreo aleatorio), lo que permitiría reconstruir correctamente las edades.

inversión (AVI), se obtiene utilizando la tasa de retorno del estudio de valoración correspondiente al periodo tarifario y el COMA proviene del mismo estudio. Por lo tanto, el VATT en el caso de ampliación tiene más riesgo que en el caso de obras nuevas. Dado que el VI de las obras de ampliación se mantiene constante por cinco periodos tarifarios, debería enfrentar menor riesgo que en el caso de las obras sujetas a VNR, en que no solo la tasa, sino que el VI también tiene riesgo²⁴.

Un primer dilema es determinar si utilizar los datos de las obras de ampliación o las de obras nuevas. Esto es importante, porque al comparar los resultados se observa que los valores AVI de adjudicación de obras nuevas recientes son muy inferiores a los valores referenciales, como se muestra en la Tabla 2. Este resultado no es explicado por unos pocos casos, porque en 49 de las 53 obras nuevas adjudicadas, el valor referencial es superior al VATT adjudicado. Los pocos casos restantes corresponden a proyectos pequeños, con VATTs solo un poco mayores que los referenciales. Es decir, las diferencias entre los valores referenciales y los resultados de las licitaciones por obras nuevas son significativas.

Por el contrario, en el caso de obras de ampliación, la Tabla 2 muestra que la diferencia es la contraria: los valores referenciales son significativamente menores que los valores adjudicados. Tampoco esto se explica por unos pocos casos. De los 129 casos adjudicados, solo 12 tienen un valor menor al referencial, y la diferencia en esos casos es de menos de un 10%. Por el contrario, los restantes casos tienen diferencias muy importantes con el valor referencial, incluso cuando hay varios participantes en las licitaciones²⁵. Por lo tanto, se puede concluir que los VI obtenidos en subastas son significativamente mayores que los VI referenciales en el caso de obras de ampliación.

Tabla 2: Diferencia entre valores referenciales y valores adjudicados.

Tipo de obras	VI o VATT Referencial	VI o VATT Adjudicado	Diferencia
Ampliación	428.9	582.3	36 %
Nuevas	154.6	81.4	-47 %

Nota: Decretos N° 418-2017, N° 294-2018, N° 198-2018, N° 422-2017, N° 231-2019, Decreto 4-2021, Decreto 13-2020.

A primera vista, parecería más probable que los valores referenciales sean menores que los de las subastas en el caso de obras nuevas. En las obras

²⁴No existe un motivo para suponer que haya alguna correlación que reduzca el riesgo entre las distintas variables que componen el VI de las obras existentes, con la tasa de descuento.

²⁵Es decir, el mayor valor persiste aún con competencia en las subastas y no se debe a la menor competencia en el caso de obras de ampliación.

nuevas, especialmente cuando se trata de líneas de transmisión, hay más riesgos en el proceso constructivo: hay riesgo de retrasos debidos a los Estudios de Impacto Ambiental (EIA), oposición local, dificultades con las servidumbres, y otros costos socioambientales. Estos son mayores que en el caso de obras de ampliación, que, en muchos casos, consisten en aumentar la capacidad de una subestación existente, hacer un cambio de conductores, o modificar pequeños tramos de líneas. Este aspecto debería elevar los costos de obras nuevas por sobre el referencial, en comparación con las obras de ampliación. Sin embargo, como se observa en la Tabla 2, los VATT de los proyectos nuevos resultan inferiores a los referenciales, en promedio.

Entonces, ¿Qué podría explicar esta diferencia en los resultados de las subastas? Una posibilidad es que las subastas de obras de ampliación sean menos competitivas que las de obras nuevas. Un examen rápido indica que hay más participantes por licitación en el caso de obras nuevas. Sin embargo, en un número importante de subastas de obras de ampliación, hay tres o más interesados. En general en esos casos los márgenes sobre el VI referencial son menores, pero esto no es siempre cierto (como uno de varios ejemplos, la diferencia con el VI referencial supera el 50% en las ampliaciones de las subestaciones Graneros y Alameda pese a recibir 5 ofertas cada una). Si se eliminaran de las subastas aquellas en las cuales solo hubo un postulante, el margen de las subastas restantes sigue siendo un elevado 26% por sobre el VI referencial. En las 60 subastas con 3 o más participantes, la diferencia se estrecha, pero sigue siendo un significativo margen de 17.5% sobre el VI referencial. En conclusión, la menor competencia en obras de ampliación es un factor que afecta los márgenes, pero no es determinante en la diferencia observada entre el valor referencial y el valor adjudicado.

En nuestra interpretación, la explicación más plausible para justificar la diferencia observada entre los resultados de las licitaciones/subastas y los valores referenciales es la diferencia en el riesgo en los retornos. El VATT de las obras nuevas es un valor fijo, que se mantiene por 5 periodos tarifarios, por lo que se parece a un bono a tasa fija durante ese período. Esto significa que la empresa puede apalancar el proyecto de forma importante, lo que es especialmente rentable cuando la tasa de interés es baja. Debido a la competencia en las licitaciones, este efecto se refleja en ofertas de VATT más agresivas. Si este hecho no está considerado en los valores referenciales de las obras nuevas, los valores de AVI alcanzados en las subastas serán menores a los AVI que se obtienen de los valores referenciales.

Con respecto a las obras de ampliación, las subastas son sobre el VI (y no sobre el VATT) de las obras. Si el costo de capital ponderado de la empresa (WACC) es mayor que el costo de capital regulado, el único mecanismo para

traspasar este mayor costo es a través de un mayor valor de oferta, lo que explica que el VI de las licitaciones de obras de ampliación sea mayor al referencial. Algo similar ocurre con la infraestructura existente, con la diferencia que en este caso no es posible traspasar el mayor costo debido a riesgo a un mayor valor de los activos. Es importante considerar, también, la similitud entre el perfil de riesgo de los ingresos en el régimen de VNR de obras existentes y el del régimen de obras de ampliación. Por todo esto, se recomienda utilizar como referencia para la valorización de activos existentes los resultados de obras de ampliación y no los resultados de obras nuevas. Esto se aborda en mayor detalle a continuación.

4.3. Por qué utilizar los resultados de licitaciones de obras de ampliación

Durante los primeros cinco períodos tarifarios, la remuneración de las inversiones en obras nuevas y en ampliaciones siguen criterios distintos del que ocurre con posterioridad, que corresponde al régimen general de VNR. Para comparar los distintos casos (obra nueva y ampliación) haremos un tratamiento separado entre lo que ocurre en los primeros 5 periodos tarifarios y la etapa posterior para una obra del mismo valor. La expresión para el valor de la obra como función de los pagos se muestra en (5).

$$V = \sum_{i=1}^{20} \frac{(VS_z)}{(1+r_z)^i} + VR \quad (5)$$

Donde V es el Valor de la obra, VS_z es el pago anual que recibe el propietario de una instalación, r_z es la tasa de descuento aplicable (debido al riesgo) durante los primeros 20 años, VR es el valor residual, considerando que a partir del sexto proceso tarifario se utiliza el método VNR de remuneración en todos los tratamientos/regímenes. Suponemos que tanto el valor del proyecto como el valor residual son los mismos para todos los tratamientos. Además, el subíndice z se refiere a los distintos tratamientos, es decir: a una obra nueva (N) o ampliación (A).

La tasa aplicable en cada caso corresponde a la tasa WACC, es decir, una tasa ponderada de la tasa de interés de endeudamiento de largo plazo r_z^l y el costo de capital de la empresa r_z^c , en que el ponderador es el factor de apalancamiento, que denominamos α , de manera que la tasa aplicable es:

$$r_z^w = \alpha r_z^l + (1 - \alpha)r_z^c, \quad z = N, A^{26} \quad (6)$$

Para simplificar, supondremos que el costo de capital es el mismo, y que ambas empresas tienen acceso a endeudamiento de largo plazo a la misma tasa²⁷. Así, la diferencia entre ambos casos es que el apalancamiento puede ser mucho mayor (más cercano a uno) si la empresa tiene asegurado un flujo de ingresos permanente por los 5 primeros períodos tarifarios que si enfrenta el riesgo de tasa de interés asociado al VI de una obra de ampliación. Dado que el endeudamiento de largo plazo tiene tasas menores que el costo de capital de las empresas, se tiene que la tasa de descuento utilizada para descontar flujos es menor en el caso de obras nuevas, es decir: $r_N^w < r_A^w$. Usando la expresión (5), esto significa que $VS_N < VS_A$. La Figura 6 ilustra el valor de VS_z para distintas tasas r_z^w . Es notable observar que la disminución de 1 punto porcentual en la tasa WACC (en la vecindad de $r_z^w = 7\%$) lleva a una disminución en torno al 8% en el valor de VS_z . Esto quiere decir que a mayor apalancamiento (que permite reducir la tasa WACC r_z^w), el mismo VI (\$100 en la Figura 6) se puede recuperar mediante pagos anuales más reducidos.

Por lo tanto, en el caso que la autoridad regulatoria no considera las diferencias en el riesgo de la remuneración entre obras nuevas y de ampliación, las valorizaciones de infraestructura calculadas a partir de los resultados de ambas subastas serán distintas, con valorizaciones significativamente menores para el caso de las obras nuevas. Recordemos, además, que los mayores riesgos constructivos en obras nuevas (dificultades por servidumbres, conflictos medioambientales, etc.) harían suponer que las valorizaciones de estas obras debieran ser mayores, lo que muestra la importancia del diferencial de tasas efectivas.

Otro punto importante a considerar son los niveles de riesgos asociados a los pagos entre los distintos regímenes (existente, ampliación, nuevo) y como esto se debiera ver reflejado en el nivel de remuneración anual. Así, comparemos el riesgo que enfrenta un activo de una obra de ampliación con el que enfrenta un activo bajo el proceso VNR. El retorno en cada período tarifario depende de dos variables aleatorias, que suponemos independientes: la valoración del activo (VI constante en el caso de obras de ampliación) y la

²⁶Por simplicidad, hemos omitido el efecto de la tasa de impuestos (T) que reduciría la tasa WACC si la empresa está apalancada, ya que un factor $(1 - T)$ debe multiplicar la tasa de interés de endeudamiento en (6).

²⁷No solo se trata, como en el ejemplo, de mayor capacidad de apalancamiento. Hay argumentos para suponer que ambas tasas (r_z^l y r_z^c) pueden ser menores en el caso de obras nuevas que en el de obras de ampliación: i) las obras nuevas a menudo son utilizadas para ingresar a un nuevo mercado, y utilizar una menor tasa de costo de capital es una inversión de entrada, ii) a menudo empresas internacionales con interés en ingresar al mercado tienen mejores condiciones en créditos de largo plazo que las empresas chilenas que usualmente realizan las obras de ampliación.

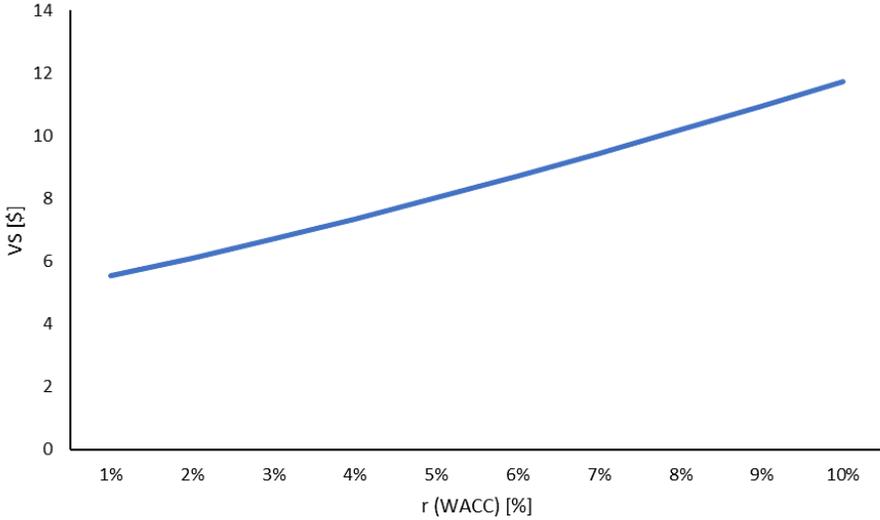


Figura 6: Valor de la anualidad en función de la tasa de descuento para un activo de $VI = \$100$ y vida útil de 20 años ($VR = 0$).

tasa de retorno en el periodo tarifario (r). Supongamos que la esperanza y la varianza en la valoración del activo bajo VNR son μ, σ^2 , respectivamente, y que la esperanza y varianza de la tasa de interés son ν, τ^2 , respectivamente. La varianza de los pagos bajo VNR es la varianza del producto, es decir: $Var(VS_V) = \sigma^2\tau^2 + \sigma^2\nu^2 + \tau^2\mu^2$ ²⁸. En el caso de obras de ampliación, no existe varianza del valor de la instalación durante los cinco primeros períodos tarifarios y solo existe varianza de la tasa de interés/retorno, por lo que la varianza de los pagos se reduce a $Var(VS_A) = \tau^2\mu^2$. Dado que los elementos de la primera suma son positivos, se tiene $VAR(VS_V) > VAR(VS_A)$. Esto muestra que hay más riesgo en el retorno bajo VNR que en el retorno asociado a las obras de ampliación para obras de igual valor real, por lo que el retorno asociado debería ser mayor. Así, los pagos anuales bajo distintos regímenes debieran respetar la siguiente desigualdad:

$$VS_V > VS_A > VS_N$$

La conclusión es que las remuneraciones asociadas a las subastas de obras de ampliación son más cercanas a las remuneraciones que deben recibir las obras bajo VNR. Los VATT obtenidos en las subastas de obras nuevas pre-

²⁸Para X e Y variables aleatorias independientes, entonces: $VAR(XY) = (VAR(X) + E(X)^2) \times (VAR(Y) + E(Y)^2) - E(X)^2 \times E(Y)^2$. Asumiendo que $VS_V \cong VNR \times r$, entonces: $Var(VS_V) = (\sigma^2 + N^2)(\tau^2 + \nu^2) - N^2\nu^2$.

sentan un perfil de riesgo menor, lo que hace que los valores ofertados sean significativamente menores.

Es un error, por lo tanto, usar la relación entre valores referenciales y resultados obtenidos en subastas de obras nuevas para ajustar los valores de VNR. Esto es importante porque, aparentemente, la CNE ajustó los valores en su estudio de valoración de activos a los valores obtenidos en las subastas de obras nuevas, por lo que los resultados de los estudios dieron valores significativamente más bajos. De hacerse un ajuste, es más apropiado considerar los resultados de las obras de ampliación, teniendo la precaución de seleccionar los valores en las licitaciones con suficientes participantes para suponer que existe competencia.

Chilquinta ha realizado el siguiente experimento para mostrar este efecto: ha utilizado la metodología de la CNE para determinar el valor de algunos activos de obras de ampliación recientes y los ha contrastado con los VI referenciales que tuvieron en su momento y el valor obtenido en la licitación. La Tabla 3 muestra estos resultados.

Tabla 3: Contraste entre VI Referencial Original, VI Adjudicado y VI según Modelo CNE, obras de ampliación. Fuente: Chilquinta.

Instalaciones	VI Referencial	VI Adjudicado	VI modelo CNE	Diferencia VIR/VICNE	Diferencia
					VIA/VICNE
Ampliación en S/E Agua Santa	9,632	11,096	8,810	-9%	-21%
Ampliación en S/E Bosquemar	2,001	2,305	1,440	-28%	-38%
Ampliación en S/E Placilla	1,814	2,090	1,539	-15%	-26%
Ampliación en S/E San Antonio	2,498	2,877	2,427	-3%	-16%
Ampliación en S/E Alto Melipilla	10,020	12,194	7,641	-24%	-37%

Como se observa en la Tabla 3, los valores del modelo CNE son incluso inferiores a los valores referenciales, sin considerar que esos valores no son realistas dado los valores adjudicados. La CNE ha sostenido que los resultados del pasado no son aplicables, pero este argumento no se aplica a los ejemplos, que corresponden a licitaciones recientes. Además, salvo el costo de los materiales y del capital, que tiene variaciones sin una dirección clara en los últimos años, los demás factores de costo son todos crecientes en el tiempo. Por lo tanto, es difícil concluir que los valores obtenidos en las obras de ampliación (que son una mejor representación de los valores VNR que los resultados de obras nuevas) hayan tenido la caída que aparece en el cuadro. Por el contrario, el ajuste realizado a los valores referenciales en la nueva metodología de la CNE parece haberse basado en los resultados de subastas de obras nuevas.

4.4. Reemplazos eficientes

Como se mencionó anteriormente, es posible demostrar que la empresa realizará un reemplazo eficiente de los activos que vayan envejeciendo, debido a que los ingresos están desacoplados de las decisiones de reemplazo, permitiendo igualar el punto donde la empresa maximiza sus utilidades con el punto de mínimo costo. No obstante, aunque la regulación actual incentiva un reemplazo eficiente, permite, a la vez, una sobre-renta para la empresa cuando la vida útil efectiva del activo supera la vida útil regulatoria. La solución evidente a este problema es reducir de forma automática la remuneración del activo antiguo una vez cumplida su vida útil regulada. Se propone remunerar con un monto algo superior al COMA, y, así, entregar una pequeña renta a la empresa, mucho menor que la que se obtiene con el mecanismo actual, que continúa pagando el costo de un activo ya pagado. Este margen permitirá que la empresa prefiera mantener el activo si es económicamente eficiente no reemplazarlo. Estos márgenes deben ser pequeños, pero deben existir para no poner en riesgo la señal de eficiencia. Esto se puede como un porcentaje del AVI, que remunera a la empresa una vez cumplida la vida útil del activo. Cuando la empresa decida reemplazar el activo, la remuneración debería volver a ser el AVI. Esta propuesta mantiene el valor del VI en el tiempo, y solo modifica el ponderador del AVI para activos que han superado su vida útil regulatoria.

El pago al activo sería igual a $\alpha \times AVI + COMA$, donde $\alpha = 100\%$ para activos que no han cumplido su vida útil y $\alpha < 100\%$ para activos cuya edad es mayor a la vida útil regulatoria. En el caso de ser necesaria una revisión de este AVI para la obra de reemplazo (por ejemplo, si la nueva obra de reemplazo es significativa), dicha obra se podría tratar de la misma forma como actualmente se tratan las ampliaciones, permitiendo actualizar el valor del VI de forma eficiente mediante licitaciones.

5. Conclusiones

Este documento analiza el mecanismo regulado para remunerar las redes de transmisión en Chile, con un particular foco en la metodología de Valor Nuevo de Reemplazo (VNR) que se utiliza para remunerar las instalaciones existentes. El análisis considera, también, la consistencia entre la regulación de la remuneración y la planificación centralizada de redes, ya que uno de los principales objetivos de la remuneración por VNR es promover inversiones eficientes en la red.

Del análisis detallado en este informe es posible concluir lo siguiente:

- El mecanismo de VNR expone a los inversionistas a una serie de riesgos producto de una valorización de los activos existentes/hundidos que varía cada 4 años.
- En términos generales, asignar estos riesgos al inversionista puede ser razonable si esto promueve la eficiencia (por ejemplo, en casos en que las empresas deciden sus expansiones).
- Dado que las decisiones sobre obras nuevas y de ampliación son mandadas por la Comisión Nacional de Energía (CNE) y que la metodología de VNR se aplica a infraestructura ya hundida, los riesgos asociados a la actualización del VNR (cada 4 años) y traspasados a la empresa para crear señales de eficiencia no tienen justificación desde un punto de vista teórico²⁹.
- Este traspaso de riesgos sólo crea ineficiencias al aumentar el costo del capital para los inversionistas.

Esto se traduce, eventualmente, en una mayor tarifa/precio para el consumidor. Por lo tanto, es recomendable, siguiendo la teoría y las mejores prácticas internacionales, minimizar estos riesgos para la empresa, ya que no son controlables por ella.

Junto con esta última recomendación, es necesario minimizar las potenciales sobre-rentas que perciben las empresas debido a que se sigue remunerando, con las mismas anualidades, a instalaciones que superan su vida útil. Para esto, se propone un mecanismo de remuneración que reduce estas sobre-rentas sin eliminar la señal de eficiencia que permite el reemplazo eficiente de las instalaciones³⁰. Esto promueve que se sigan utilizando las instalaciones mientras los costos de operarlas sean menores que el costo producto de fallas y de empeoramiento de la calidad de suministro.

²⁹De hecho, en la opinión de los autores, la valorización por VNR se heredó erróneamente de la anterior regulación. Esto es un error porque el actual marco de planificación y de remuneración no son compatibles (en el antiguo marco de la subtransmisión, por ejemplo, efectivamente las empresas decidían la expansión de sus redes y la planificación realizada durante el proceso tarifario no era vinculante).

³⁰Las decisiones de reemplazo (realizadas una vez que los activos han cumplido su vida útil) permanecen en manos de las empresas, lo cual podría estimarse como una justificación para mantener el mecanismo de VNR. En opinión de los autores, esta justificación es insuficiente, dado que el costo del riesgo originado por el mecanismo de remuneración es significativamente superior a sus beneficios. Además, proponemos una manera eficiente de tratar el reemplazo de activos antiguos simple y consistente con los principios regulatorios (ver propuesta en Sección 4.3).

Es importante reducir en forma paralela tanto las sobre-rentas como los riesgos excesivos asociados a valoración por VNR. En la opinión de los autores, uno de los problemas que genera la existencia de sobre-rentas es que tienden a deslegitimar el procedimiento regulatorio, lo que podría llevar al regulador a adoptar prácticas que implican expropiación regulatoria.

Como reemplazo al VNR, se propone una valorización fija (no cambiante en el tiempo) para los activos existentes. Esta valoración debería reducirse al cumplirse la vida útil contable y hasta que sea reemplazado por un activo nuevo. Esto requiere llevar una contabilidad de la base de activos y sus antigüedades, la que ya se posee en forma parcial. Nuestra propuesta se detalla en este documento.

Finalmente, el reemplazo del VNR por una valorización fija requiere definir los Valores de Inversión (VI) iniciales para las instalaciones existentes. Este documento demuestra (tanto teóricamente como en base observaciones empíricas de los resultados de las licitaciones) cómo usar los resultados de licitaciones de obras para cumplir este objetivo. Observamos que, incluso en subastas con varios participantes, el valor de las obras de ampliación es mayor que el valor de referencia, y que lo contrario ocurre en el caso de obras nuevas. La explicación es que el perfil de riesgo de la remuneración es mucho menor en el caso de obras nuevas (que se parecen a un bono durante el periodo de invariabilidad) que en el caso de obras de ampliación³¹. El perfil de riesgos bajo VNR es más parecido al de las remuneraciones de obras de ampliación. Por ello, sugerimos que es inapropiado ajustar los valores de inversión usando valores observados en subastas de obras nuevas y que, si se desea ajustar los valores de referencia, es más apropiado hacerlo usando los resultados de las licitaciones de las obras de ampliación. Otra alternativa para fijar los valores de inversión iniciales de instalaciones existentes es utilizar valores de VNR provenientes de fijaciones tarifarias anteriores que hayan sido aceptados por las partes. Con esto, se consigue el objetivo de minimizar los riesgos de valoración, dado que estos crean ineficiencias y elevan los costos del negocio y, así eventualmente, la tarifa final.

³¹El VATT de las obras nuevas es un valor fijo, que se mantiene por 5 periodos tarifarios, por lo que se parece a un bono a tasa fija durante ese período. Esto significa que la empresa puede apalancar el proyecto de forma importante, lo que es especialmente rentable cuando la tasa de interés es baja. Esto justifica que las ofertas alcanzadas en las subastas de obras nuevas sean menores a los valores de referencia, en contraste con lo que ocurre en el caso de obras de ampliación, en que es el valor de inversión, pero no la remuneración, el que se mantiene constante.

Referencias

- [1] S. K. Abeygunawardane, P. Jirutitijaroen, y H. Xu. Adaptive maintenance policies for aging devices using a markov decision process. *IEEE Transactions on Power Systems*, 28(3):3194–3203, 2013.
- [2] D. Alvarado, R. Moreno, M. E. Orchard, y D. S. Kirschen. Cost-benefit analysis of maintenance plans: Case study of the power system of a large industrial facility. *IEEE Transactions on Power Systems*, 2022.
- [3] H. Averch y L. L. Johnson. Behavior of the firm under regulatory constraint. *The American Economic Review*, 52(5):1052–1069, 1962.
- [4] A. Damodaran. Betas by sector (us), nov 2022. Disponible en http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/Betas.html.
- [5] E. Engel, R. Fischer, y A. Galetovic. Licitación de carreteras en Chile. *Estudios públicos*, (61), 1996.
- [6] E. M. Engel, R. D. Fischer, y A. Galetovic. Least-present-value-of-revenue auctions and highway franchising. *Journal of political economy*, 109(5):993–1020, 2001.
- [7] R. D. Fischer y P. Serra. *Efectos de la privatización de servicios públicos en Chile: Casos sanitario, electricidad y telecomunicaciones*. Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Facultad . . . , 2004.
- [8] T. Gómez. Monopoly regulation. En *Regulation of the Power Sector*, páginas 151–198. Springer, 2013.
- [9] W. W. Hogan. Contract networks for electric power transmission. *Journal of regulatory economics*, 4(3):211–242, 1992.
- [10] T. Irwin. Government guarantees: Allocating and valuing risk in privately financed infrastructure project-washington. *DC: International Bank for Reconstruction and Development/World Bank*, 2007.
- [11] J. D. Jenkins y I. J. Pérez-Arriaga. Improved regulatory approaches for the remuneration of electricity distribution utilities with high penetrations of distributed energy resources. *The Energy Journal*, 38(3), 2017.
- [12] P. L. Joskow. Incentive regulation in theory and practice: electricity distribution and transmission networks. *Economic regulation and its reform: What have we learned?*, páginas 291–344, 2014.

- [13] J.-J. Laffont y J. Tirole. A theory of incentives in regulation and procurement, 1993.
- [14] C. Matamala, R. Moreno, y E. Sauma. The value of network investment coordination to reduce environmental externalities when integrating renewables: Case on the chilean transmission network. *Energy Policy*, 126:251–263, 2019.
- [15] R. Moreno, B. Bezerra, H. Rudnick, C. Suazo-Martinez, M. Carvalho, A. Navarro, C. Silva, y G. Strbac. Distribution network rate making in latin america: An evolving landscape. *IEEE Power and Energy Magazine*, 18(3):33–48, 2020.
- [16] R. Moreno, M. Panteli, P. Mancarella, H. Rudnick, T. Lagos, A. Navarro, F. Ordonez, y J. C. Araneda. From reliability to resilience: Planning the grid against the extremes. *IEEE Power and Energy Magazine*, 18(4):41–53, 2020.
- [17] Ofgem. *Handbook for implementing the RIIO model*. Ofgem London, 2010. Disponible en https://www.ofgem.gov.uk/sites/default/files/docs/2010/10/riio_handbook_0.pdf.
- [18] D. Papadaskalopoulos, Y. Fan, A. D. Paola, R. Moreno, G. Strbac, y D. Angeli. Game-theoretic modeling of merchant transmission investments. En *Transmission Network Investment in Liberalized Power Markets*, páginas 381–414. Springer, 2020.
- [19] L.-M. Perrin. Mapping power and utilities regulation in europe. *EY Global Power & Utilities Center, EYGM Limited*, 2013. Disponible en: https://erranet.org/wp-content/uploads/2018/04/Mapping_regulation_in_Europe.pdf.
- [20] R. Serrano, M. R. Carvalho, J. C. Araneda, O. Alamos, L. Barroso, D. Bayma, R. Ferreira, y R. Moreno. Fighting against wildfires in power systems: Lessons and resilient practices from the chilean and brazilian experiences. *IEEE Power and Energy Magazine*, 20(1):38–51, 2022.

Instituto Sistemas Complejos de Ingeniería

INSTITUTO SISTEMAS COMPLEJOS DE INGENIERÍA

ISCI
Instituto Sistemas
Complejos de Ingeniería

El Instituto Sistemas Complejos de Ingeniería (ISCI) trabaja en un continuo que va desde el desarrollo de metodologías de frontera hasta aplicaciones innovativas en Sistemas Complejos de Ingeniería, en los cuales interactúan infraestructura, comportamiento humano y gestión.

GRUPOS DE INVESTIGACIÓN

- Recursos Naturales: Minería, Forestal y Acuicultura
- Transporte
- Localización y Logística
- Consumer Analytics
- Smartcities
- Data Science
- Organización Industrial
- Energía

PATROCINA:



INSTITUCIONES ALBERGANTES:



ANEXO DEL ESTADIO
PUEBLO Y MARTIN GARCÍA
UNIVERSIDAD DE CHILE



INSTITUCIÓN ASOCIADA:



contacto@isci.cl
www.isci.cl

Programas de Postgrado y Postítulos DII



Doctorado en Sistemas de Ingeniería



Un Programa Interdisciplinario
en Investigación Avanzada

Formamos especialistas con una sólida base tecnológica y un conocimiento profundo de las herramientas que permiten modelar, entender y optimizar sistemas complejos en que interactúan elementos físicos de comportamiento humano.

www.dsiuchile.cl
doctorado@sistemasdeingenieria.cl
56 229784017
Beauchef 851
Torre Oriente 2° Piso
Oficina 204

7
AÑOS

Universidad Acreditada
Diciembre de Postgrado
Gestión Institucional
Diciembre de Postgrado
Vinculación con el medio
Nov 2018 - Nov 2025

4
AÑOS

Doctorado Acreditado
Agosto 2016 a
Agosto 2020
CNA-Chile


Comisión Nacional
de Acreditación
CNA-Chile

Programas de Postgrado y Postítulos DII

**Con la Beca Mujer MGO
tienes un descuento del
100% del arancel**

Todas las alumnas aceptadas
postulan a la Beca

Postula para Otoño 2022



MAGÍSTER EN GESTIÓN DE OPERACIONES

El Magíster busca formar graduados de excelencia en investigación de operaciones, quienes podrán enfrentar problemas complejos en gestión de operaciones, integrando herramientas matemáticas, económicas y tecnológicas.

+200

Graduados en empresas destacadas y doctorados de clase mundial

 **Cuerpo docente de excelencia**
 **Infraestructura de primer nivel**



"Soy extranjera y así como ésta beca definió mi presente, podría definir el futuro de muchas mujeres, que sólo necesitan que se les dé una oportunidad".

Maria José Jiménez

"Incentivar el ingreso de mujeres a postgrados en ingeniería, donde estamos subrepresentadas tanto entre alumnos como profesores, aumentará nuestra participación en esta área laboral y académica".

Gabriela Alfaro



"La Beca Mujer es una muy buena iniciativa para motivar a más a mujeres a entrar al MGO y lo he notado, porque después de que entré, hay más mujeres que han decidido postular".

Javiera Morales

"La Beca Mujer influyó en mi decisión de entrar al Magíster en Gestión de Operaciones. Siento que esta beca es importante porque me hace sentirme valorada en un área en donde aún hay pocas mujeres y con esta posibilidad se puede lograr mayor equidad".

Daniela Valdovinos



Para aumentar la participación de mujeres en STEM y particularmente en el MGO

**Beca Mujer
MGO**

100% descuento de arancel

Infórmate en www.mgo.uchile.cl



POSTULA AL MAGÍSTER EN GESTIÓN DE OPERACIONES
www.mgo.uchile.cl

7
AÑOS

Universidad Fernández
Análisis de Operaciones
Gestión de Operaciones
Ingeniería de Operaciones
Ingeniería de Sistemas

8
AÑOS

Magíster Asesorado
Especialización en
Logística y
Operaciones


Comisión Nacional
de Acreditación
CNA-Chile

Programas de Postgrado y Postítulos DII



MAGÍSTER EN GESTIÓN Y POLÍTICA PÚBLICAS

PROGRAMA LÍDER EN CALIDAD Y COMPROMISO PÚBLICO

Referente en la formación de profesionales que cuentan con los conocimientos, las competencias y las habilidades para liderar los procesos de diagnóstico, comprensión, resolución y priorización de problemas públicos, vinculando el conocimiento con las necesidades de los países y resguardando los criterios de eficiencia, eficacia, impacto y legitimidad ciudadana de las acciones públicas.

REFERENTE EN AMÉRICA LATINA Y EL CARIBE EN LA
FORMACIÓN EN GESTIÓN Y POLÍTICAS PÚBLICAS

COMPOSICIÓN INGRESO 1995 - 2020

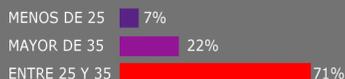
79% CON BECAS
(AGCID, SERVIR, PRONABEC,
ANID, F. KONRAD ADENAUER,
MGPP, ENTRE OTRAS)



DIVERSIDAD PROFESIONAL



EDAD AL INGRESO (AÑOS)



EQUIDAD DE GÉNERO



Síguenos en redes sociales

MGPP_UChile
@MGPPUChile
@MGPP_DII

Más información

www.mgpp.cl
mgpp@dii.uchile.cl
+56 9 3580 2171
+56 9 5610 5485

Programas de Postgrado y Postítulos DII



MBA UNIVERSIDAD DE CHILE
INGENIERÍA INDUSTRIAL

PRIMEROS EN GESTIÓN

TALLERES DE HABILIDADES DIRECTIVAS
QUE TE PREPARAN PARA CARGOS DE ALTO LIDERAZGO

ENFOQUE DEL MANAGEMENT SCIENCE PARA
TOMAR MEJORES DECISIONES DE GESTIÓN

MBA EXECUTIVE

24 o 20 meses

Viernes (tarde) y sábado cada 2 semanas

MBA PART TIME

20 o 16 meses

3 a 4 veces a la semana (tardes)

Av. República 701, Santiago, Chile.
+56 2 2978 4002

8
AÑOS

RECONOCIDO POR COMISIÓN
NACIONAL DE CALIFICACIÓN
DE GRADUADOS
Mayo 2017 - Agosto 2018

Comisión Nacional
de Acreditación
CNAAC



MBA UChile
Ingeniería Industrial | Universidad de Chile

Programas de Postgrado y Postítulos DII



MAGCEA
INGENIERÍA INDUSTRIAL
UNIVERSIDAD DE CHILE




MAGÍSTER EN ECONOMÍA APLICADA

OBJETIVO DEL MAGÍSTER

El Magíster en Economía Aplicada forma profesionales y académicos de gran capacidad analítica y sólida base en economía.

3 LÍNEAS DE ESPECIALIZACIÓN

- Microeconomía
- Finanzas y Macroeconomía
- Métodos Empíricos

CUERPO ACADÉMICO

Publicaciones en las principales revistas académicas y conferencias internacionales
Investigación Internacional
Iniciativas que favorecen la investigación

Para promover la participación de mujeres en economía

Beca Mujer en economía y ciencias

100% descuento de arancel (2 cupos máxima por año)



+200
GRADUADOS

PHD en Universidades de clase mundial	37%
Sector privado y en la Banca	44%
Organismos reguladores y sector público	19%

7 AÑOS

Universidad Acreditada por el Consejo Nacional de Investigación Científica y Tecnológica (CNA-Chile) desde mayo 2018 - Nov 2023

8 AÑOS

Magíster Acreditado por el Consejo Nacional de Investigación Científica y Tecnológica (CNA-Chile)



Comisión Nacional de Acreditación CNA-Chile

POSTULA AL MAGÍSTER EN ECONOMÍA APLICADA

www.magcea-uchile.cl

Desarrolla una mirada estratégica de tu organización

y lidera proyectos de innovación tecnológica en procesos de negocios

15 AÑOS DE EXPERIENCIA



MBE
MASTER IN BUSINESS ENGINEERING
INGENIERÍA INDUSTRIAL | UNIVERSIDAD DE CHILE

7 AÑOS

Universidad Acreditada por el Consejo Nacional de Investigación Científica y Tecnológica (CNA-Chile) desde mayo 2018 - Nov 2023

3 AÑOS

Magíster Acreditado por el Consejo Nacional de Investigación Científica y Tecnológica (CNA-Chile) desde mayo 2018 - Nov 2023



Comisión Nacional de Acreditación CNA-Chile



CNEP CONSEJO NACIONAL DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA

Duración: 3 semestres académicos

www.mbe.cl

Contacto: mbe@dii.uchile.cl | +56 22 978 4835

Programas de Postgrado y Postítulos DII



MBA

INDUSTRIA MINERA
INGENIERÍA INDUSTRIAL
UNIVERSIDAD DE CHILE



MAGÍSTER EN GESTIÓN Y DIRECCIÓN DE EMPRESAS PARA LA INDUSTRIA MINERA

FORMANDO LÍDERES PARA LOS DESAFÍOS DE LA NUEVA MINERÍA

mbamin.cl

inicio de clases, abril 2022



Datos clave

Inicio
31 marzo 2023

Horario
Part. time en Chile, full time extranjero

Duración
9 meses en la Universidad de Chile
+ 8 meses intercambio o doble grado en Canadá, Reino Unido o Australia
+ Global Experience Trip en China
+ 2 meses Proyecto Grado (a distancia)



Requisitos

3 años experiencia laboral desde egreso de pre-grado

Licenciatura o título profesional universitario con duración oficial mínima de 8 semestres

Inglés avanzado
Ver puntajes exigidos en pruebas estandarizadas de dominio del idioma (TOEFL o IELTS) y GMAT (sólo para Cranfield), según la escuela de negocios elegida



Fase Internacional

Intercambio
UBC, Canadá

Doble grado
Cranfield, Reino Unido
Macquarie, Australia

Global Experience Trip: China
(Incluido en el costo del programa)



Inversión

Aplica según universidad socia.

Intercambio:
USD 26.500 - UBC, Canadá*

Doble grado:
USD 33.500 - Cranfield, Reino Unido*
USD 48.500 - Macquarie, Australia*

*Valores incorporan Beca UBC y Beca Cranfield

Programas de Postgrado y Postítulos DII

EDUCACIÓN EJECUTIVA




UNIVERSIDAD
DE CHILE

EDUCACIÓN EJECUTIVA
INGENIERÍA INDUSTRIAL

EXCELENCIA PARA IMPULSAR TU CARRERA
Diplomados | Cursos de Especialización
Programas Corporativos

www.eeuchile.cl

