



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

DISEÑO E IMPLEMENTACIÓN DE HERRAMIENTA DE ASIGNACIÓN DIARIA DE CONTENEDORES MINIMIZANDO EL COSTO LOGÍSTICO PARA UNA EMPRESA FORESTAL.

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

JUAN FRANCISCO JAVIER PÉREZ URREA

PROFESOR GUÍA:
PATRICIO CONCA KEHL

MIEMBROS DE LA COMISIÓN:
RICARDO SAN MARTÍN ZURITA
CHARLES THRIVES CORTÉS-MONROY

SANTIAGO DE CHILE
2020

RESUMEN EJECUTIVO

En el presente informe se presenta el trabajo que se desarrollará durante la memoria de título en la empresa Arauco, compañía forestal chilena dedicada a la fabricación de pulpa de celulosa y derivados como madera aserrada y paneles.

El 95% de la producción de la empresa se exporta hacia el extranjero en contenedores y carga a granel, cumpliendo un rol muy importante el nivel de servicio que se les entrega a los clientes para así fidelizarlos cada día más. El problema es que actualmente esto no se está logrando y los pedidos no están llegando a tiempo a los principales clientes de la compañía, como es el caso del mercado de Estados Unidos, dónde el volumen entregado a tiempo a los clientes ha disminuido notablemente entre los meses de mayo y julio, desde un 74% de nivel de servicio a 59% respectivamente.

Los datos operacionales que son imputados en SAP contienen errores que, debido a su gran volumen, son difíciles de identificar y corregir, impidiendo que herramientas como simuladores y optimizadores funcionen adecuadamente.

El volumen de exportación de Arauco se traduce en cerca de 150 mil contenedores que tienen que viajar en 550 naves hacia 242 puertos distintos. En Chile la empresa tiene 26 plantas productivas, y embarca en 4 puertos distintos. Elegir en que nave embarcar una entrega a un cliente es una tarea difícil, pues debe tener en consideración variables de producción, contratos con las navieras, fechas de arribo, tiempo de viaje, restricciones aduaneras, capacidades portuarias, costos, nivel de cumplimiento con el cliente, entre otras. Incurrir en una mala elección implica que la carga llegue tarde al cliente, lo cual tiene un costo indirecto para Arauco como lo es la pérdida de una venta o fidelización de un cliente si es que la carga llega atrasada, y costos directos como los costos de transporte, por ejemplo, embarcando en una nave más cara, o costos extraportuarios, cuando se llega adelantado al puerto de destino, por almacenar carga por un periodo de tiempo mayor al contratado.

El trabajo propuesto tiene como objetivo general desarrollar una herramienta de asignación de contenedores a buques centrada en el usuario, con el fin de alcanzar el nivel de servicio meta de Arauco al menor costo posible. La propuesta a entregar consiste en la implementación de un sistema de procesamiento de datos, el cual sea capaz de reportar y corregir errores en los datos, generar las transformaciones necesarias y facilitar la comunicación con los tomadores de decisión, además un modelo de optimización de asignación de contenedores a buques, cuya función objetivo es alcanzar el nivel de servicio meta de la compañía al mínimo costo posible para el área de Logística.

Los resultados muestran que es posible ahorrar hasta \$600.000 USD al año en costo logístico manteniendo el mismo nivel de servicio actual. Por otro lado, se observa que es posible aumentar el nivel de servicio por sobre el 95% incurriendo en un mayor gasto.

TABLA DE CONTENIDO

| | | |
|---------|--------------------------------------------------------------|----|
| I. | ANTECEDENTES GENERALES..... | 1 |
| | I.I INTRODUCCIÓN | 1 |
| | I.II VOLUMENES PRODUCTIVOS | 3 |
| | I.III Posicionamiento internacional..... | 4 |
| II.III | DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN..... | 6 |
| III.III | OBJETIVOS | 12 |
| IV. | MARCO CONCEPTUAL..... | 13 |
| V. | METODOLOGIA | 14 |
| | V.I DIAGNOSTICO SITUACIÓN ACTUAL..... | 14 |
| | V.II DEFINIR EL PROYECTO | 14 |
| | V.III MODELO DE DATOS | 14 |
| | V.IV PREDICCIÓN PLANTA PRODUCTIVA | 14 |
| | V.V MODELO DE OPTIMIZACIÓN..... | 15 |
| | V.VI VALIDACIÓN DE LA HERRAMIENTA | 15 |
| VI. | ALCANCES | 16 |
| VII. | DIAGNOSTICO SITUACIÓN ACTUAL | 17 |
| VIII. | MODELO DE DATOS | 19 |
| | VIII.I SELECCIÓN HERRAMIENTA PROCESAMIENTO DE DATOS..... | 19 |
| | VIII.II DEPURACIÓN DE DATOS | 20 |
| IX. | PREDICCIÓN PLANTA PRODUCTIVA | 22 |
| | IX.I EL PROCESO KDD | 22 |
| | IX.II MÉTODOS DE MINERÍA DE DATOS | 23 |
| | IX.II RESULTADOS PREDICCIÓN PLANTA PRODUCTIVA | 30 |
| X. | MODELO DE OPTIMIZACIÓN | 32 |
| | X.I MODELO DE MINIMIZACIÓN DE FLETE..... | 35 |
| | X.II MODELO MAXIMIZACIÓN NIVEL DE SERVICIO | 38 |
| | X.III RESTRICCIONES IMPLÍCITAS, CONJUNTO DE PARES | 40 |
| | X.IV IMPLEMENTACIÓN Y MANTENCIÓN DE LA HERRAMIENTA | 42 |
| XI. | RESULTADOS | 44 |
| | XI.I AHORRO COSTO LOGÍSTICO ANUAL | 44 |
| | XI.II FLETE VS NIVEL DE SERVICIO | 45 |
| | XI.III ANÁLISIS FUNCIÓN OBJETIVO DEL NIVEL DE SERVICIO | 47 |
| | XI.V REASIGNACIONES | 49 |
| | XII.V ANÁLISIS DE SENSIBILIDAD | 51 |
| XII. | CONCLUSIONES Y RECOMENDACIONES | 53 |
| XIII. | BIBLIOGRAFÍA..... | 55 |

I. ANTECEDENTES GENERALES

I.1 INTRODUCCIÓN

Arauco es una empresa forestal chilena dedicada a la fabricación de pulpa de celulosa y derivados de la madera como madera aserrada y paneles, cuenta con plantaciones de pino y eucalipto, las cuales, según la Corporación Chilena de Madera CORMA, se encuentran entre las de mayor crecimiento en el mundo. La empresa pertenece al grupo económico de Anacleto Angellini, Antar Chile. Es la principal empresa forestal de Latinoamérica, tiene plantas productivas en Chile, Argentina, Uruguay, Brasil y Estados Unidos, y, además, cuenta con oficinas en Australia, Colombia, Perú, México, Holanda, Japón, Corea y Canadá, en total cuenta con más de 15 mil trabajadores y más de 3.500 clientes en los 5 continentes. En el territorio chileno, tiene 18 plantas productivas, 8 aserraderos y 3 plantas químicas, posee más de 1.1 millones de hectáreas de patrimonio forestal y el año 2018, reportó ventas por 5.953 millones de dólares, el monto más grande en la historia de la compañía, consolidando la estabilidad de la compañía.

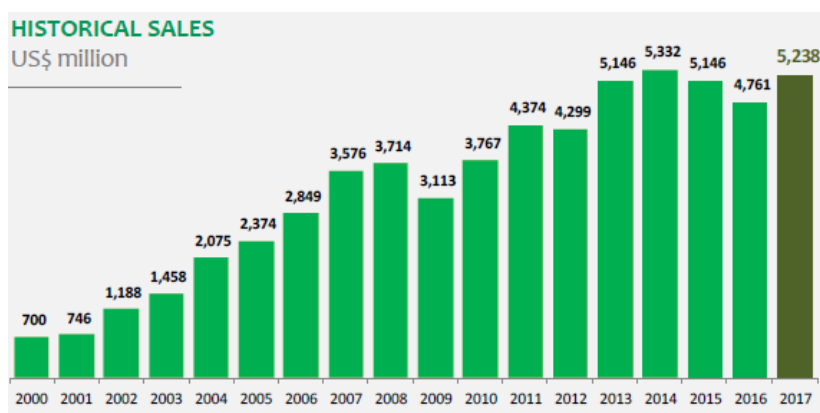


Gráfico 1: Nivel de ventas históricas de Arauco.

Fuente: Presentación corporativa Antar 2018.

Su misión consiste en promover un manejo forestal que sea ambientalmente sustentable, socialmente beneficioso y económicamente viable en los bosques de todo el mundo, lo que se complementa con su visión de contribuir a mejorar la vida de las personas, desarrollando productos forestales para los desafíos de un mundo sostenible.

Según memoria anual de Arauco de diciembre del 2017, Arauco se ubica como el segundo mayor productor de celulosa del mercado mundial y segundos en la producción de madera aserrada y paneles. A nivel nacional, Arauco es líder en exportaciones de productos forestales en términos de ingresos por ventas y número uno en cuando a volumen total exportado, dando a la compañía un importante poder de negociación de fletes marítimos y terrestres.

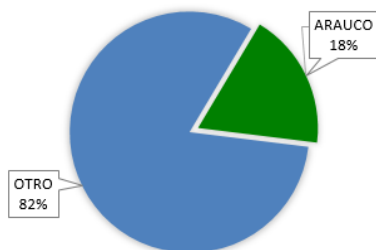


Gráfico 2: Porcentaje de contenedores totales exportados en 2016.
Fuente: Aduana Chilena 2017.

El sector forestal se posiciona como uno de los pilares fundamentales de la economía de Chile, con una participación del 2,5% del PIB nacional. La industria destaca por su aporte social y económico, generando cerca de 120.000 empleos directos.

Se presenta a continuación el organigrama de Servicios Logísticos Arauco S.A (SLASA), área donde se desarrollará el trabajo de memoria, cuyo objetivo es llevar el producto terminado al cliente final.

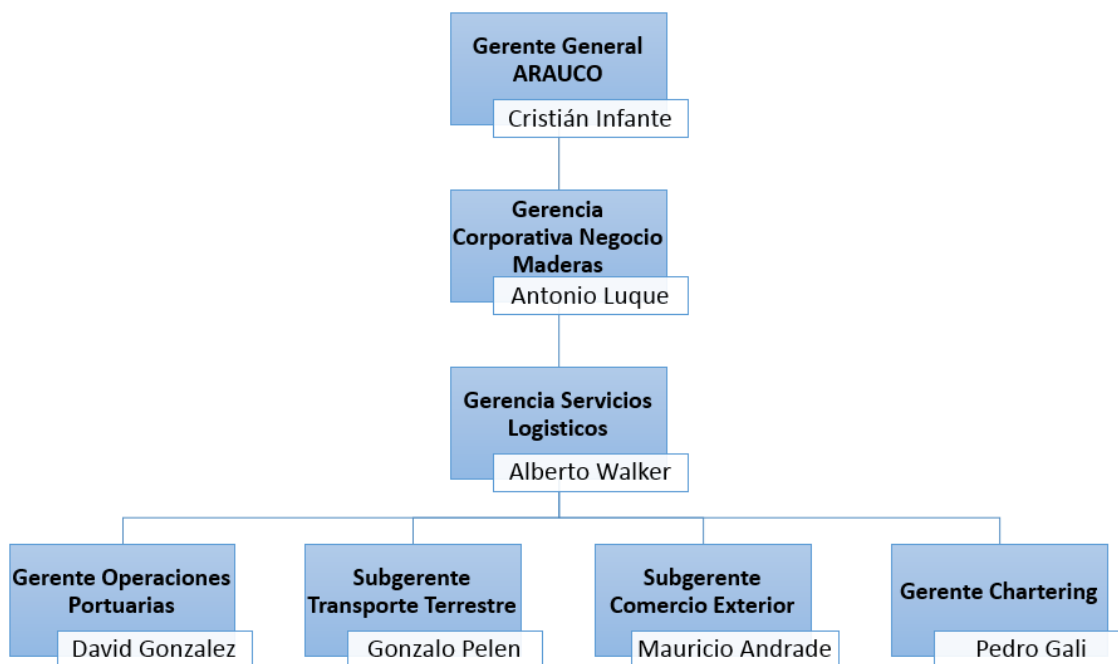


Imagen N°1: Organigrama Gerencial SLASA.
Fuente: Intranet Arauco mayo de 2019.

I.II VOLUMENES PRODUCTIVOS

La celulosa es uno de los commodities de mayor presencia en los productos de uso cotidiano en la vida de las personas. La celulosa es utilizada para la manufactura de papel de impresión y escritura, papel higiénico y sanitario, cartulinas y material de empaque. Según el gráfico N°3, Latinoamérica posee la mayor capacidad de producción mundial de celulosa, según cifras del año 2015.

DISTRIBUCIÓN DE LA CAPACIDAD DE PRODUCCIÓN MUNDIAL DE CELULOSA DE MERCADO AÑO 2017
(70,4 millones de toneladas)

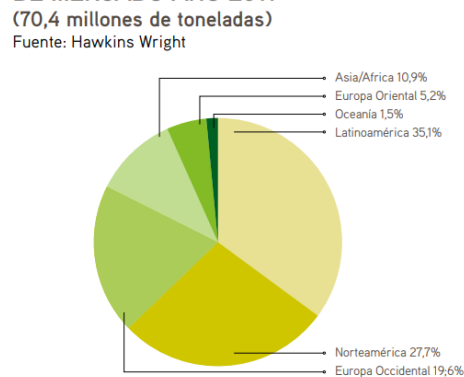


Gráfico 3: Capacidad de producción celulosa.

Fuente: Memoria Anual ARAUCO 2017.

Arauco se enfoca principalmente en la comercialización de dos productos, Celulosa y Maderas. Dentro de la Celulosa, existen 3 tipos de este material: Pulpa blanqueada de Eucalyptus Globulus (EKP), Pulpa blanqueada de Pinus Radiata (BKP) y Pulpa no blanqueada de Pinus Radiata (UKP).

| Producto | 2015 | 2016 | 2017 |
|---------------------------------|------------------|------------------|------------------|
| Pino Blanqueada (BSK=BKPR+BKPT) | 1.422.139 | 1.531.089 | 1.441.910 |
| Eucalipto Blanqueada (BEK) | 1.528.236 | 1.526.333 | 1.662.565 |
| Pino Cruda (UKP) | 454.304 | 429.947 | 435.051 |
| Fluff | 198.460 | 208.129 | 155.153 |
| Total | 3.603.139 | 3.695.498 | 3.694.679 |

Tabla 1: Producción de celulosa por tipo de producto (adt).

Fuente: Memoria Anual ARAUCO 2017.

Dentro de Maderas, existen más de diez familias de productos, en donde destacan las familias de Madera Aserrada Seca, Madera Aserrada Verde, MDF, Molduras, Plywood, Aglomerados, entre otras.

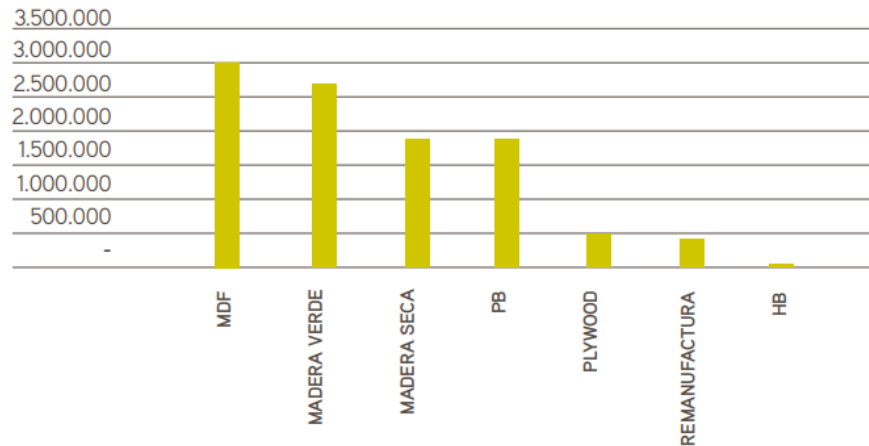


Gráfico 4: Producción de Madera por tipo de producto en miles de m3 en 2017.
Fuente: Memoria Anual ARAUCO 2017.

La empresa exporta el 95% de su producción a los 5 continentes desde los puertos de coronel, San Vicente, Lirquén y San Antonio, estos traslados se realizan de dos formas: a través de buques portacontenedores, que transportan carga en contenedores estandarizados de 20 y 40 pies; y a través de buques Break Bulk, que transportan carga empaquetada, que se carga y descarga en forma individual. El 75% del volumen que se exporta, se traslada en Contenedores y el 25% restante se realiza en naves Break Bulk.

De la carga que se traslada en Contenedor, un 51% corresponde a Madera, mientras que el 49% restante corresponde a Celulosa. Además, el 86% del volumen de Celulosa se exporta al continente asiático, siendo China el principal cliente con un 81% del volumen de Celulosa exportado. Por otra parte, el 33% del volumen de Madera es exportado a Estados Unidos y el 27% se exporta al continente asiático.

I.III Posicionamiento internacional

Arauco, en cuanto a capacidad productiva se refiere, está dentro de las empresas forestales más grandes de mundo. Ocupa el segundo lugar en producción de celulosa y de paneles, productos con mayor valor en el rubro forestal. En América Latina sus principales competidores son Fibria, empresa cuyas operaciones se centran en Brasil, y la empresa de origen chileno CMPC, con operaciones en varios países de Sudamérica y en México.

El 90% de la capacidad productiva de Arauco se encuentra en Sudamérica, el otro 10% en Norteamérica y Europa. En el mercado europeo, sus principales competidores son empresas de origen ruso, con las cuales Arauco no puede competir debido a desventajas de costo logístico y tiempos de viaje.

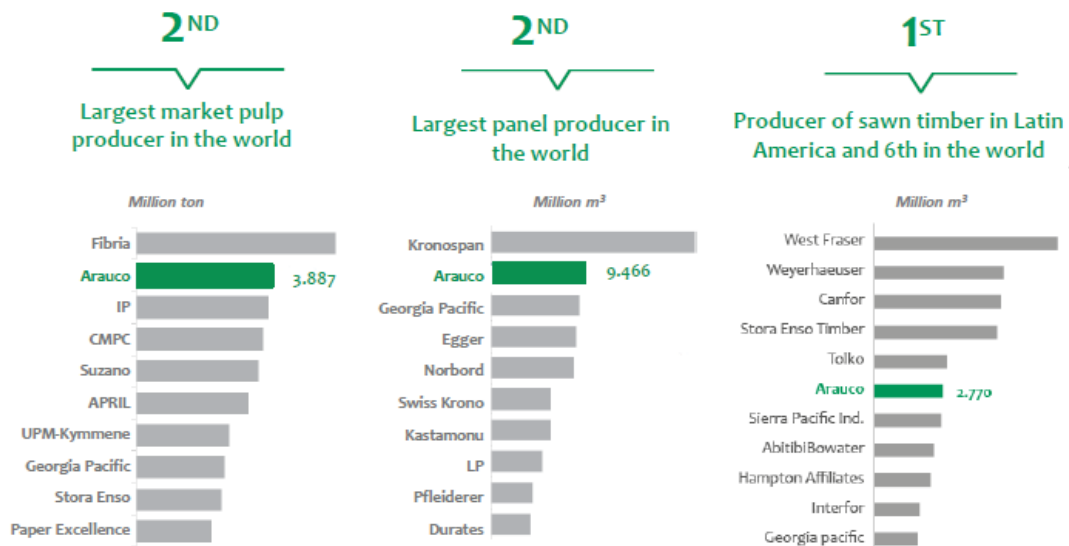


Imagen N°2: Posicionamiento de mercado forestal Arauco.

Fuente: Memoria Anual ARAUCO enero 2019.

II.III DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN

Servicios Logísticos Arauco, es una gerencia de la empresa que se encarga de prestar servicios a las distintas áreas del negocio, buscando contribuir a la cadena de entrega de productos terminados a su cliente final. Para esto, cuenta con cuatro áreas que se encargan de diferentes partes de la cadena. El área de transporte terrestre, de comercio exterior, el área de operaciones portuarias y el área de Chartering o de exportación. Esta última se encarga de asignar una entrega a un buque para que un pedido pueda llegar al cliente en el extranjero, cerciorándose de que este pedido llegue dentro del tiempo pactado al menor costo posible para Arauco.

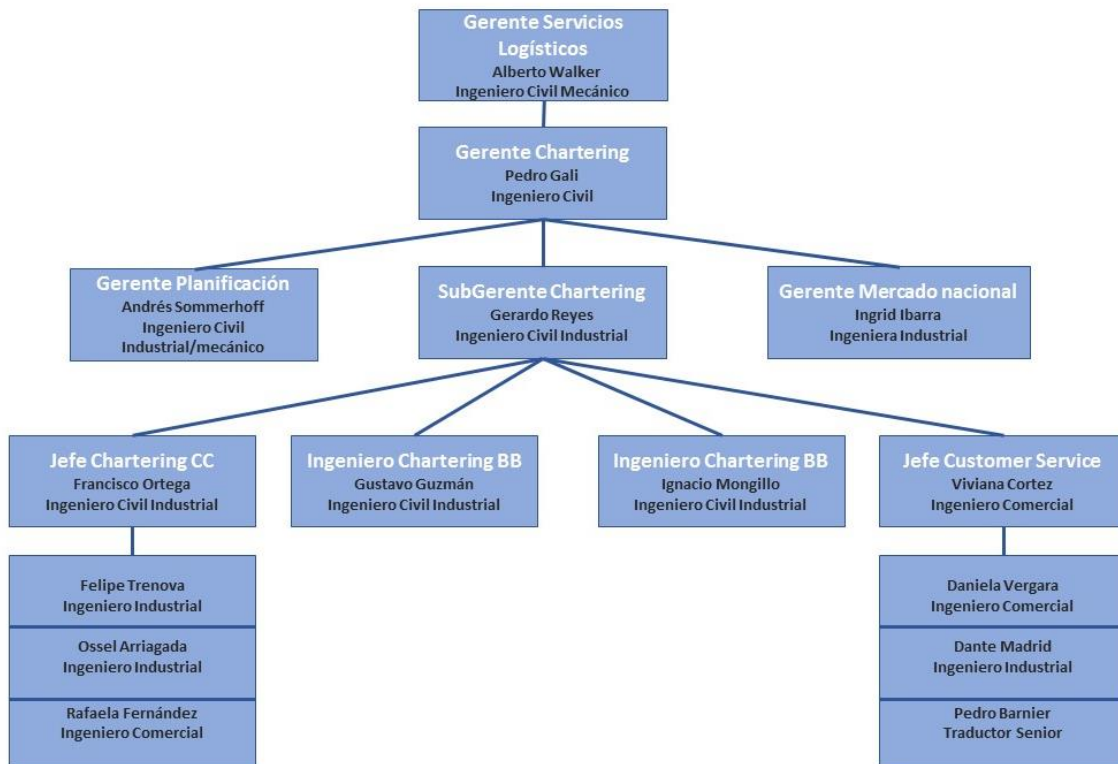


Imagen N°2: Organigrama área de exportación (Chartering).

Fuente: Intranet ARAUCO mayo 2019.

Como se mencionó anteriormente, la empresa Arauco exporta más del 95% de su producción, lo que al año se traduce en cerca de 150 mil contenedores que tienen que viajar en 550 naves hacia 242 puertos distintos, para que el producto pueda llegar al cliente. Para la empresa, una Entrega es un pedido u orden que realiza un cliente, la cual involucra desde 1 hasta n contenedores dentro de la misma entrega, dependiendo del volumen pedido. A una entrega, se le hace seguimiento desde que el cliente realiza el pedido hasta su llegada al cliente, para posteriormente, medir el nivel del servicio otorgado. Este proceso se puede observar en la Imagen N°3.

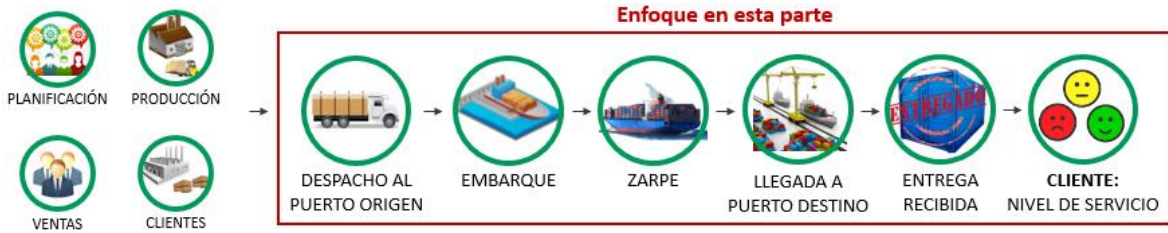


Imagen N°3: Proceso completo de una entrega.

El principal problema que se presenta actualmente es el incumplimiento de los tiempos de entrega comprometidos con los clientes. Para evidenciar esto, se analizó el mercado de Estados Unidos, uno de los mercados más importantes para la empresa, dado que representa un tercio de las ventas de madera. Este mercado es muy exigente ya que la competencia directa se encuentra dentro del mismo país, por lo tanto, los clientes exigen cumplir el mismo plazo de entrega que les ofrece una empresa productora local, porque en el caso contrario, prefieren comprar en ese mismo mercado, lo que implica que, para las empresas de este mercado, el cumplimiento de los plazos de entrega sea muy importante.

Una forma utilizada actualmente para medir el nivel de servicio es a través de Fill-Rate. El Fill-Rate (FR) es un indicador que mide la cantidad que se le entrega al cliente con respecto a lo que solicitó, pero la forma utilizada por la empresa añade la medición de atrasos en las entregas a este indicador. Según los mercados, existe una holgura de tiempo en que se puede adelantar o atrasar la llegada de una entrega, siendo ± 10 días en el caso de Estados Unidos, por lo tanto, una entrega que llega posterior a diez días de la fecha pactada se considera Atraso y una entrega que llega anterior a diez días de la fecha pactada se considera Adelanto, ambos casos son considerados como no entregado a tiempo. El gráfico N°4 muestra el nivel de servicio de Arauco con sus clientes según el mes de llegada de la entrega, desde enero hasta agosto de 2018 son datos reales, septiembre y octubre son proyecciones. Hay que considerar que el nivel de servicio meta promedio de Arauco es del 80%.

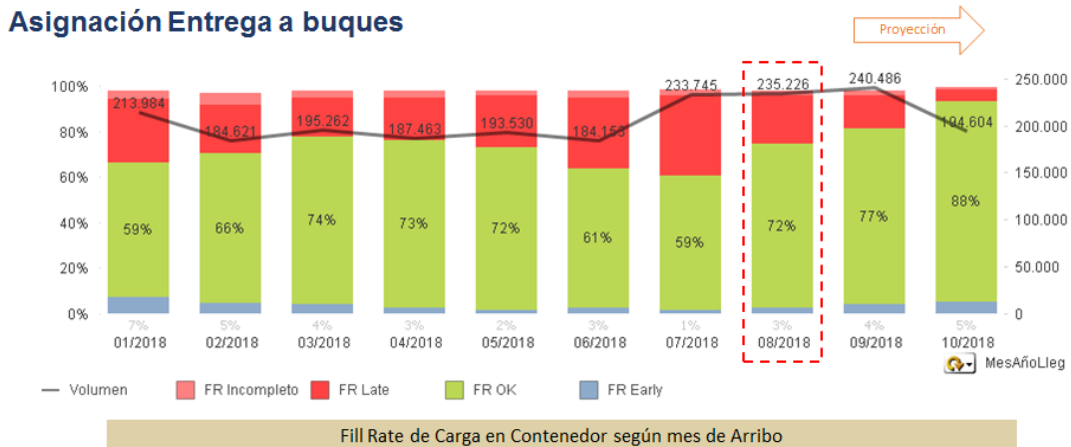


Gráfico N°4: Nivel de servicio Arauco USA. Fuente: Presentación de Coordinación SLASA.

Fuente: Presentación de coordinación SLASA Agosto de 2018.

La misión de las operaciones de SLASA es alcanzar las metas de nivel de servicio al cliente de Arauco al menor costo posible, pero elegir en que buque embarcar una entrega a un cliente es una tarea difícil, pues debe tener en consideración variables de producción, contratos con las navieras, fechas de arribo, tiempo de viaje, restricciones aduaneras, capacidades portuarias, entre otras. Incurrir en una mala elección implica que la carga llegue tarde al cliente, lo cual tiene un costo indirecto para Arauco como lo es la pérdida de una venta o fidelización de un cliente si es que la carga llega atrasada, y costos directos como los costos de transporte, por ejemplo, embarcando en una nave más cara, o costos extraportuarios, cuando se llega adelantado al puerto de destino, por almacenar carga por un periodo de tiempo mayor al contratado.

Cada una de las partes de la cadena de abastecimiento de Arauco es muy compleja por sí sola, provocando que el área de producción y logística trabajen como módulos independientes y no como un todo, es decir, logística decide en que puerto embarcara una entrega sin conocer la planta productiva o almacén de dicha entrega, y producción decide la planta productiva sin saber el puerto de embarque de la entrega.

Lo anterior revela una deficiencia entre los costos de transporte navieros, que abarca desde el puerto de origen a puerto de destino, y costos terrestres desde la planta productiva a los puertos de origen, en dónde estos últimos pueden ser incluso más caros que el primero.

La tabla 2 muestra un análisis de costos terrestres en el sur de Chile, para el producto Madera Verde en el año 2018, el cual equivale al 7% del volumen total exportado en contenedores para el mismo año, se compara el costo de asignar las entregas/contenedores de este producto al puerto de origen más lejano versus el costo de asignarlo al más cercano. Se observa que, para el producto analizado, el posible beneficio anual es de 395.380 dólares.

| | |
|---------------------------------------------|----------------------------|
| Producto | Madera Verde |
| Puerto | LQN, SVT, CNL |
| Plantas | ARAUCO, VALIVIA.... |
| Año | 2018 |
| Costo de asignar peor Puerto-Planta | X + 395.380 \$USD |
| Costo de asignar Mejor Puerto-Planta | X \$USD |
| Diferencia | 395.380 \$USD |

Tabla 2: Análisis de costos terrestres.

Fuente: Elaboración propia a partir de exportaciones totales de Arauco del año 2018.

Por el lado de logística, es natural pensar en una herramienta computacional que ayude en la difícil tarea de asignación. En la historia del área se han desarrollado Optimizadores estratégicos y operacionales, el último desarrollo en el ámbito operacional fracasó, pues la mala calidad de los datos impedía su correcto funcionamiento, provocando que quedara obsoleto y sin uso 2 meses después de su puesta en marcha.

Cabe destacar que, en el optimizador operacional logístico, no se incorporaron los costos terrestres, pues por temas operacionales y logísticos con las navieras, las entregas se asignan de 2 a 3 semanas antes de la fecha de arribo de la nave, y para esa fecha, se desconoce la planta productiva del 40% de las entregas, lo cual hace imposible conocer a priori el costo terrestre desde la planta productiva al puerto de embarque. Un tipo de producto se puede producir en una o más plantas, tal como se muestra en la tabla 3, así que, si se pretende incorporar los costos terrestres en las decisiones del área logística, se tendrá que predecir la planta productiva del 40% de las entregas.

| Producto | Plantas productivas |
|--------------|---------------------|
| Apariencias | 4 |
| Blanks | 5 |
| Madera Verde | 6 |
| Madera Seca | 6 |
| Cel BKP | 3 |
| Cel EKP | 2 |
| Cel UKP | 2 |
| HB | 1 |
| Molduras | 7 |
| Plywood | 2 |
| Laminados | 1 |
| Pallets | 1 |

Tabla 3: Diversidad plantas productivas por producto.

Fuente: Elaboración propia a partir de datos de producción del año 2018.

II.I OPTIMIZADOR DE ASIGNACIÓN OBSOLETO

El ERP usado en Arauco es SAP, es aquí donde se almacenan los datos operacionales de la empresa como estados productivos, ventas y transacciones de los clientes, movimiento de carga, transportes navieros disponibles entre otros. Son alrededor de mil empleados los que tienen acceso a este sistema y, como todo ser humano, se equivocan imputando datos. Si bien un error en los datos puede no tener gran impacto en la operación de la empresa, sí puede provocar que sistemas complejos de análisis de datos no funcionen correctamente, pues esperan cierto formato de *input*, por ejemplo, en una operación matemática se espera un número, pero se recibe una letra.



✘ Datos Erróneos difíciles de identificar y modificar por los usuarios.

Imagen N°4: Herramienta de asignación entrega/contenedores a buque.

La herramienta consta de tres partes, la obtención de los datos, el optimizador en sí y la interfaz con los resultados. La sábana de datos necesaria para alimentar el optimizador proviene de 5 fuentes:

- Datos de ventas, como las nuevas entregas a despachar, el puerto de destino y el tipo de contrato (Despacho a puerto o a un domicilio en tierra por camión o tren).
- Datos de las navieras, como los tiempos de tránsito y las fechas de arribo de las naves.
- Datos de producción, como la fecha de término de producción, el avance de la entrega y el actual centro de acopio.
- Datos internos de Chartering, como los fletes navieros por Ruta, puerto de destino. Y tipo de contenedor, y los términos del contrato con las navieras.
- Desde SAP se recogen las naves con su capacidad actual para madera y celulosa, los puertos que visita esa nave y la carga actual asignada a dicha nave hasta la fecha.

En total, son 500 columnas y 600.000 filas de datos que deben ser modificados para alimentar el optimizador de asignación, el cual, dadas las condiciones dinámicas de la empresa, debe ejecutarse de una a tres veces al día, en un tiempo máximo de diez minutos.

La última herramienta desarrollada se implementó en octubre del año 2016, la cual, según el jefe de Chartering, cuando el sistema tenía algún error en los datos se tardaban al menos una hora en encontrarlo y corregirlo, y cuando se lograba ejecutar no podían confiar en los resultados, pues desconocían la calidad de los datos con que se alimentó el optimizador. Señala, además, que los cambios más relevantes en la asignación deben realizarse antes de las diez de la mañana por lo tanto ejecutaban el sistema por solo diez minutos, logrando un GAP de optimalidad del 20%, este GAP, según lo que informaba el sistema, equivale a \$200.000 dólares aproximadamente.

Desde el año 2016, nuevos contratos navieros, condiciones portuarias y nuevas rutas marítimas han cambiado las restricciones operacionales de logística, obligando a rehacer el modelo de optimización antiguo y buscar alguna forma de que, con las nuevas restricciones, mantener o mejorar el tiempo de resolución.

Para calcular el costo de oportunidad que existe debido a la falta de una herramienta de optimización que ayude en la asignación de las entregas, se tomaron todas las entregas que ya tenían nave asignada, cuyo embarque se realizaría hasta en 3 semanas más, y se asignaron a la nave más barata, sin preocuparse del nivel de servicio, respetando solo la capacidad de las naves, se llega a que el costo logístico total se puede reducir en \$366.000 USD.

Análogamente, si se asignan las mismas entregas a la mejor nave, considerando solo el nivel de servicio estimado, sin tener en cuenta el costo, se llega a que el nivel de servicio estimado se puede mejorar un 12%. De esta forma, anualizando el ahorro logístico, se puede decir que el costo potencial de no tener una herramienta de optimización es una combinación lineal entre 3 millones de dólares y un 12% de nivel de servicio al cliente.

III.III OBJETIVOS

III.I Objetivo General

Desarrollar e implementar herramienta de asignación de contenedores a buques a nivel operacional, centrada en los tomadores de decisión, con el fin de alcanzar el mejor nivel de servicio al menor costo posible.

III.II Objetivos Específicos

1. Realizar un levantamiento del proceso de asignación actual de contenedor a buque.
2. Investigación y elección de las herramientas analíticas a utilizar.
3. Identificar fuentes de errores en los datos y reportarlo a los responsables.
4. Construir sistema de extracción, transformación y generación de los datos necesarios para el optimizador de asignación.
5. Identificar potencial beneficio de incorporar costos terrestres.
6. Construir herramienta de predicción de planta productiva.
7. Desarrollar modelo de optimización y simular.
8. Análisis de resultados.
9. Incorporar herramienta en el uso diario del equipo de Chartering.

IV. MARCO CONCEPTUAL

A continuación, se presenta el marco teórico que se utiliza como base para esta memoria:

Gestión del cambio centrada en el usuario: Gestión del cambio es un enfoque estructurado para gestionar los aspectos de cambio relacionado con las personas y la organización, esto implica realizar estudios formales de impacto de cambio, desarrollar planes de acción individuales, mejorar las comunicaciones y proporcionar capacitación para contrarrestar la resistencia. En la etapa de implementación, se obtendrá constante retroalimentación de los usuarios, los cuales entregarán información de uso en cada etapa del diseño.

Modelamiento Lineal: La programación lineal (LP, también llamada optimización lineal) es un método para lograr el mejor resultado (como el beneficio máximo o el costo más bajo) en un modelo matemático cuyos requisitos están representados por relaciones lineales. La programación lineal es un caso especial de programación matemática (también conocida como optimización matemática).

Más formalmente, la programación lineal es una técnica para la optimización de una función objetivo lineal, sujeta a restricciones de igualdad y desigualdad lineales. Su región factible es un politopo convexo, que es un conjunto definido como la intersección de muchos medios espacios, cada uno de los cuales está definido por una desigualdad lineal. Su función objetivo es una función afín (lineal) de valor real definida en este poliedro. Un algoritmo de programación lineal encuentra un punto en el poliedro donde esta función tiene el valor más pequeño (o más grande) si existe dicho punto.

Modelo de Predicción: Son representaciones simplificadas de la realidad por medio de un conjunto de hipótesis, las cuales son usadas para la explicación de algunos patrones de comportamiento que se observan en el mundo real a través de datos. En la literatura se describe un sinnúmero de modelos de predicción, el desempeño del modelo utilizado no se puede conocer a priori, sino que se evaluará una vez realizado un análisis exploratorio de los datos y comparando el nivel de precisión que entregue cada modelo.

Data WareHousing: Todos los datos necesarios se obtendrán de la herramienta de gestión SAP y almacenes de datos locales del área de logística. Los datos serán elegidos, transformados, almacenados y almacenados utilizando herramienta KNIME. La herramienta KNIME proporciona una interfaz de programación interactiva con el usuario, facilitando el reconocimiento de errores y su corrección.

V. METODOLOGIA

V.I DIAGNOSTICO SITUACIÓN ACTUAL

En esta etapa se busca entender la situación actual de la logística de Arauco, para efectos de comprensión del problema, representar sus operaciones con el resto de la cadena de suministro de Arauco, levantar los distintos niveles de decisiones del área de Chartering, entrevistar a las personas interesadas en el proyecto y comprender sus dolores.

El objetivo es evaluar la complejidad de la herramienta a desarrollar, cuáles son sus alcances y limitaciones. También hacer una medición de los costos más relevantes del área, la volatilidad de los tiempos de traslado y su impacto en la cadena logística, para tener una variable de que ayude a definir las decisiones a ser modeladas.

V.II DEFINIR EL PROYECTO

A través de este proceso se busca definir el problema a abordar y el objetivo general que aseguren el cumplimiento del proyecto. Esta etapa está enfocada a elegir cuál de las decisiones del área de logística son los que más contribuyen al objetivo estratégico de la compañía. Las decisiones son variables medibles de desempeño del trabajo logístico, estas tienen que ser con reducción de costos y mejorar la calidad de servicio al cliente.

Esto se desarrolla investigando la estrategia de negocios de Arauco, luego, como este trabajo está enfocado en el nivel de servicio al cliente, es necesario averiguar la visión estratégica de la empresa orientada al cliente. Finalmente se definen los objetivos de la herramienta de asignación el cual busca mejorar la calidad de los datos, reducir el tiempo utilizado en la toma de decisiones, minimizar los costos logísticos y cumplir con las metas de nivel de servicio de Arauco.

V.III MODELO DE DATOS

Se establece las arquitecturas tecnológicas y de sistema que proveerán apoyo al análisis de datos y al modelo de optimización y su ejecución. Las decisiones que se tomarán a través del modelo y el diseño centrado en el usuario serán las que entreguen los requerimientos funcionales y no funcionales para un correcto diseño de la herramienta de asignación y su implementación.

V.IV PREDICCIÓN PLANTA PRODUCTIVA

En esta etapa se pretende predecir en qué planta productiva será procesada la entrega, a través de un modelo de minería de datos, esto permitirá conocer el costo de transporte hasta el puerto de embarque de las entregas. La minería de datos es una de las técnicas encaminadas al descubrimiento de la información que se encuentra contenida en los datos. Analiza patrones, comportamientos, tendencias, asociaciones y características que

existen en los datos, con el objetivo de ordenarlos para extraer de manera automatizada la inteligencia contenida en los datos a través de métodos supervisados y no supervisados. Dentro de las metodologías más utilizadas de minería de datos esta la KDD, del acrónimo inglés Knowledge Discovery in Database, el cual es un proceso de descubrir conocimiento a través de patrones válidos y potencialmente útiles a través de los datos. Uno de los procesos dentro de KDD es el usuario, que ya es él quien determina el dominio de la aplicación y los datos a utilizar.

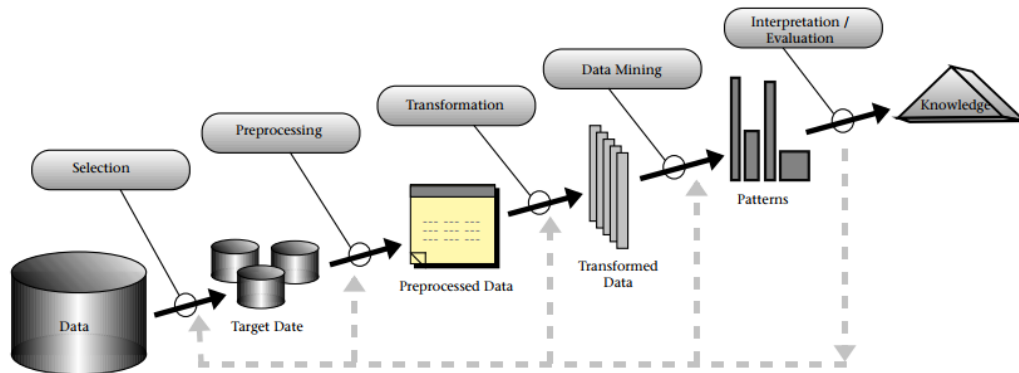


Imagen N°5: Metodología KDD, Knowledge Discovery in Database

V.V MODELO DE OPTIMIZACIÓN

Teniendo a mano la bibliografía adecuada, se procede a escoger los elementos que pueden ser utilizados en el problema en particular. Este punto es muy importante, dado que se establecen líneas de análisis que se abordarán a lo largo del trabajo de memoria. Además, se debe decidir qué tipo de modelamiento se utilizará, algunas de las posibles opciones son: Programación Lineal, Heurísticas o una combinación de ambas. En este punto se definen las variables decisión del problema de asignación, las cuales forman parte del entregable una vez terminada la memoria. Se deben formular todas las restricciones involucradas en el problema de asignación de contenedores a buques, e incluir variables auxiliares que permitan identificar cuando una restricción no puede ser cumplida.

V.VI VALIDACIÓN DE LA HERRAMIENTA

En esta etapa se busca validar la solución del optimizador con los tomadores de decisión y la forma en que esta se muestra. La solución estará validada si cumple con todas las restricciones y criterios operacionales de la asignación de una entrega a un buque, para ello se necesitará la constante retroalimentación del área de Chartering.

Además, se capacitará a los usuarios de la herramienta en cuanto al código, pues es necesario que ante algún problema sepan cómo solucionarlo a la brevedad, ya que, de no ser así, el proyecto estaría en riesgo de quedar obsoleto en poco tiempo.

VI. ALCANCES

El trabajo se realiza en la empresa Arauco y constitución S.A, ubicada en Las Condes, Santiago. Específicamente en la gerencia de Servicios Logísticos Arauco en el área de planificación logística.

Toda herramienta desarrollada durante el proyecto es para apoyar la toma de decisiones del área de Chartering de la empresa, específicamente en las decisiones que tiene que ver con el transporte de productos terminados en Chile hasta el cliente final en contenedores.

Se busca que el proyecto a realizar sea lo más útil para la empresa y también perdurable en el tiempo. La propuesta a entregar consiste en la implementación de un sistema de procesamiento de datos, el cual sea capaz de reportar y corregir errores en los datos, generar las transformaciones necesarias y facilitar la comunicación con los tomadores de decisión, además un modelo de optimización de asignación de contenedores a buques, cuya función objetivo es alcanzar el nivel de servicio meta de la compañía al mínimo costo posible para el área de Logística.

El trabajo no considera el desarrollo de una interfaz personalizada para observar los resultados, pero se realizará un análisis en detalle de datos, los cuales serán suficientes para ser utilizados en una futura interfaz a implementar.

VII. DIAGNOSTICO SITUACIÓN ACTUAL

Antes de comenzar con el desarrollo de la herramienta, se describe el proceso completo de una entrega para lograr un mayor entendimiento de la situación actual. El proceso completo de una entrega consta de tres partes, la primera parte consiste en el proceso de venta, la segunda parte es el proceso de producción, y por último el proceso de entrega del pedido.

Creación de un pedido

El proceso de venta de una entrega comienza cuando un cliente realiza un pedido y termina cuando el área de ventas ingresa el pedido al sistema. Por su parte, el área de ventas tiene que verificar con el área de planificación la disponibilidad para vender el producto que solicita el cliente, es decir, verificar la factibilidad del pedido, cantidad disponible para vender y fechas en que se puede entregar este producto al cliente. El área de planificación es la encargada de comunicar esta información al área de ventas en base a los tiempos establecidos por el área de producción y el área de logística. Ventas propone una orden de compra al cliente donde se detalla la cantidad de producto, precio y promesa de embarque. La promesa de embarque puede ser una fecha de zarpe de la entrega desde puerto chileno o una fecha de llegada al lugar donde el cliente lo solicite.

Una vez se ingresa el pedido al sistema, el área de producción se encarga disponer el producto solicitado para que puede ser enviado al cliente, en forma paralela, el área de Chartering se encarga de asignar el pedido a una de las naves de las navieras disponibles para realizar el envío.

Fechas relevantes

Dentro de este proceso existen tres fechas muy importantes, la Fecha Preferente de Asignación de Nave, en adelante FPAN, la cual contempla el tiempo de termino de producción del pedido, el tiempo de transporte terrestre, consolidación del contenedor y aduanas. La FPAN puede cambiar si producción lo requiere, agregando días cuando hay retraso productivo o restando días en el caso de adelantos, estos cambios de FPAN pueden implicar un cambio de nave. Otra fecha importante es la Fecha estimada de arribo de la Nave, en adelante ETA por sus ciclas en inglés Estimated Time of Arrival, que corresponde a la fecha que llegara la nave al puerto de origen, esta fecha la informa cada naviera para sus respectivas naves. Posterior a la ETA, la nave permanece dos días en el puerto, para permitir la carga y descarga de contenedores, y luego zarpa. Por último, se encuentra la Fecha Requerida por el Cliente, en adelante FRC, la cual indica cuando la entrega debe llegar al cliente.

Proceso de asignación

Para entregas cuyo acuerdo con el cliente es una fecha de embarque, Chartering debe buscar una nave que llegue al puerto de origen (San Antonio, Liquén, Coronel o San Vicente) entre dos días antes de la FPAN y 6 días después de esta. Para entregas con promesa de fecha de llegada a destino, Chartering debe buscar una nave con ETD

(Estimated Time to Destiny) entre 10 días antes o 10 días después de la FRC. La ETD se calcula como ETA más 2 días (estimación del Zarpe de la nave), más el tiempo de tránsito en agua y el tiempo de tránsito en tierra estimado.

El tiempo de tránsito en tierra, se refiere al tiempo entre que el buque arriba en puerto de destino hasta que llega al lugar de destino, es cual es un estimado que depende de la distancia al puerto y el tipo de transporte, que puede ser camión o tren. El tiempo de tránsito en agua al puerto de destino es informado constantemente por las navieras y cargado a las bases de datos locales al área de Chartering. En el gráfico N°5 se muestra el cumplimiento de este tiempo de tránsito, si observamos la cantidad de contenedores que llegan atrasados o adelantados debido a un retraso de la nave en el agua, estos representan el 12% del total. Se puede decir que el tiempo de viaje presenta una buena estimación.

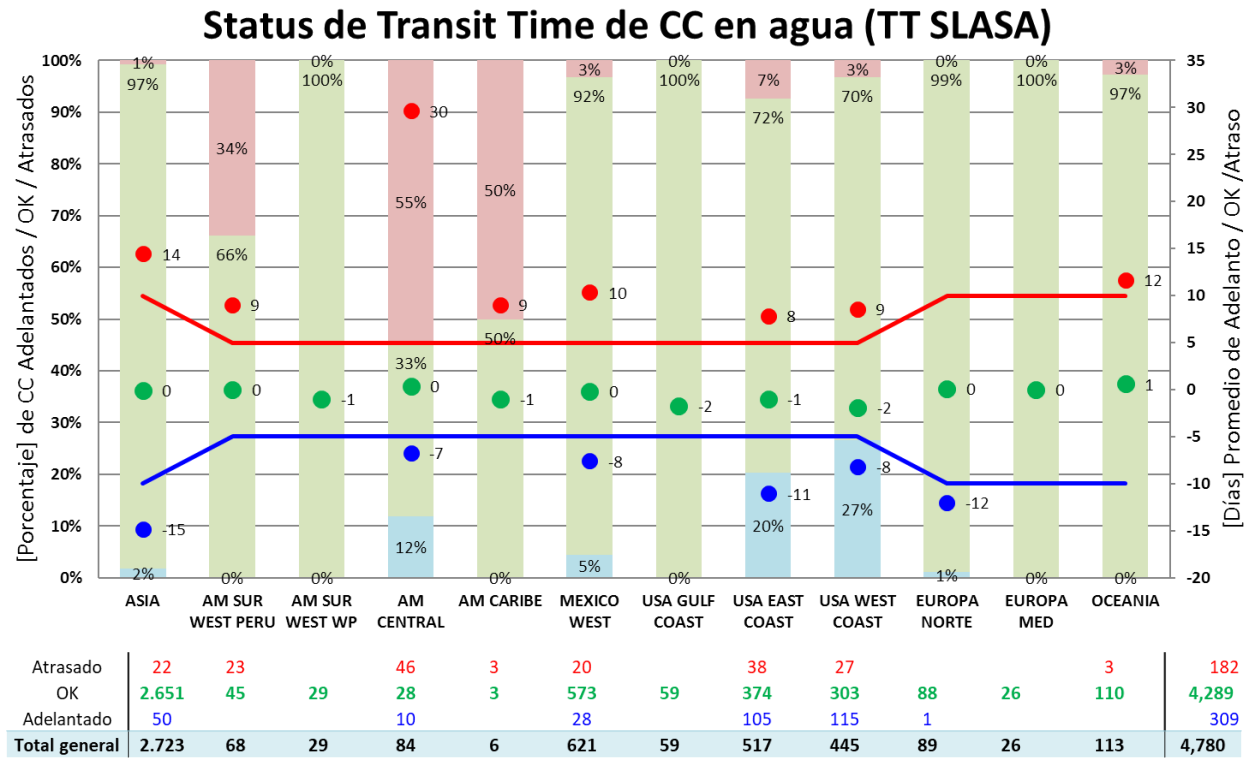


Gráfico N°5: Transit Time Agua por mercado.
Fuente: Presentación de Coordinación SLASA, marzo 2019.

El proceso de asignación de entregas a naves se realiza desde 30 días antes de la FPAN hasta 7 días antes, dado que entre más cerca de la FPAN los contenedores se consolidan y se facturan con la naviera asignada. El proceso se realiza de lunes a viernes entre las ocho y las diez de la mañana y consiste en una serie de decisiones y revisiones para cada una de las entregas en este rango. Entre las revisiones más usuales están los cambios de FPAN de las entregas, los cambios de ETA y tiempos de tránsito, cambios de puerto de destino y nuevas entregas a asignar.

VIII. MODELO DE DATOS

VIII.1 SELECCIÓN HERRAMIENTA PROCESAMIENTO DE DATOS

KNIME (o Konstanz Information Miner) es una plataforma de minería de datos que permite el desarrollo de modelos en un entorno visual, tal como se observa en la imagen N°6. La plataforma es de código abierto y gratis a nivel empresarial, es fácil de escalar e intuitiva de aprender. KNIME permite manipular datos, analizarlos, transformarlos y realizar flujos de trabajo auto comentados en base a nodos que se conectan entre sí. Lo que hace atractivo y potente a KNIME es la gran cantidad nodos preconstruidos, y dado que está montado en Eclipse, la comunidad a nivel mundial puede aportar con desarrollos de nuevos nodos. Hay nodos para hacer casi cualquier cosa, si hay alguna tarea que no se pueda hacer nativamente, hay nodos que se conectan con módulos de Python, Java, R, Weka y Matlab.

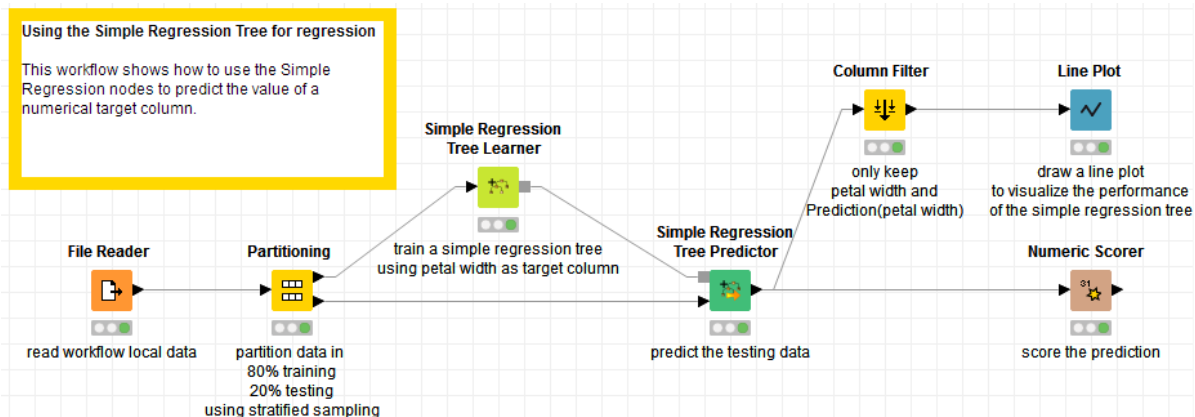


Imagen N°6: Ejemplo desarrollo en plataforma KNIME.

La aplicación es ideal para personas poco entendidas con el mundo de la programación. Hacer prototipos es realmente fácil y rápido, tanto que no es necesario un conocimiento previo o muy avanzado para diseñar una ruta en KNIME. Si bien el área de chartering está compuesta de ingenieros, es su rutina diaria nunca han tenido que programar algún código, si se les presenta un desarrollo escrito únicamente en java o Python corre el peligro de quedar obsoleta rápidamente, incluso ser rechazada de inmediato. Que los usuarios entiendan por completo el desarrollo de la herramienta es primordial, si ocurre algún error en el proceso es necesario que sea corregido ágilmente por los mismos usuarios.

Para corregir cualquier tipo de error, lo primero es identificarlo, KNIME también ayuda a los usuarios en esta labor. Tal como se muestra en la imagen N°7, los nodos tienen un estado que muestran la respuesta del procesamiento. Se puede decir entonces que el semáforo clásico (Rojo-Amarillo-Verde) entrega la información general del estado de un nodo, pero que bien podría tener algunos detalles (% de ejecución o tipo de error

encontrado) para el paso entre el amarillo y el verde. Además de esto, KNIME también podría informar alertas, advirtiendo sobre efectos colaterales que se deben considerar, generalmente relacionadas al manejo de archivos de lectura y escritura, como, por ejemplo, el riesgo de sobrescribir un documento o que el documento no exista. Además, los resultados del procesamiento de cada nodo pueden ser consultados en cualquier momento; en la opción de visualización, ubicada al final del menú desplegable.

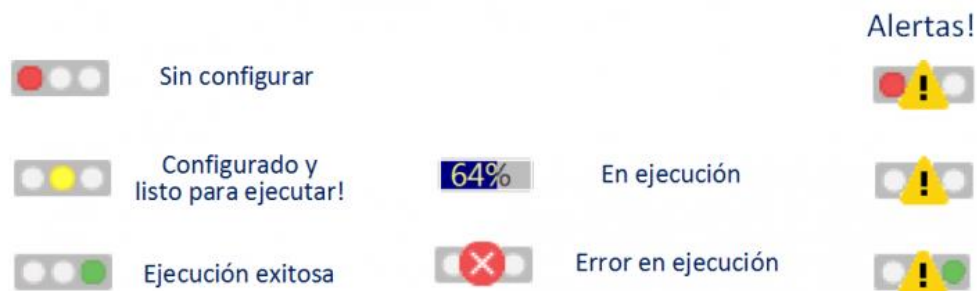


Imagen N°7: Estados posibles de un nodo en KNIME.

Fuente: Elaboración Propia.

VIII.II DEPURACIÓN DE DATOS

La empresa cuenta con toda la información referente a una entrega dentro de su plataforma SAP en ambientes separados. Existe un ambiente SAP para las entregas que ya cuentan con una nave asignada, un ambiente para entregas sin nave, otro para el estado productivo de las entregas y un cuarto donde vive el conjunto de naves disponibles dentro de los próximos dos meses. La información asociada a logística se encuentra en un archivo Excel, denominado datos maestros, el cual se actualiza constantemente en el tiempo, perdiendo información histórica de este archivo. En los datos maestros se puede encontrar información específica asociada a las naves como, los tiempos de tránsito, rutas marítimas, itinerarios, puertos de trasbordo, fletes y capacidad de contenedores.

Actualmente existen procedimientos y rutinas que entregan información importante para la gestión, pero no cuentan con toda la información necesaria para tomar decisiones óptimas. Los datos maestros no siempre cuentan con información completa y los datos de SAP no siempre son correctos debido a la intervención humana en el llenado de datos. Para generar una herramienta de asignación es necesario contar con información confiable, para ello es necesario depurar los datos, con ello se busca tener en cada momento información completa y sin errores.

En la tabla 4 se muestran los errores más comunes de los cuales el modelo de datos se hace cargo, cada solución fue acordada en conjunto con el área de Chartering. Las soluciones planteadas tienen el objetivo de reflejar la realidad de forma aproximada con tal de minimizar los errores en el proceso de optimización.

| Tipo de Error | Descripción | Solución |
|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Entrega sin Contenedores o con una cantidad errónea.</i> | La cantidad de contenedores que necesita una entrega tiene origen en SAP, la cual es proporcional al volumen del pedido. | Cuando la relación volumen/contenedores se pierde, se recalcula con un factor de consolidación aproximado según tipo de producto (EJ: 30 m3/CC para Celulosa). |
| <i>Entrega sin tipo de Contenedor</i> | El tipo de contenedor que usará una entrega puede ser de 20 o 40 pies de largo. El tipo de contenedores es un campo en SAP y define el flete marítimo que se pagará. | El 90% de las entregas se embarca en contenedores de 40 pies, así que por defecto se corregirán los campos vacíos por un tipo de 40 pies. |
| <i>Entregas de un mismo grupo de entregas con diferente FRC O FPAN</i> | Un grupo de entregas debe ser embarcado en la misma nave. Pero por construcción, si estas entregas tienen distintas fechas, es muy probable que no se asocien con las mismas naves, haciendo infactible al modelo. | Operativamente, cuando esto ocurre, el área de planificación actualiza la FPAN y FRC, replicando la fecha más lejana al resto de las entregas del grupo. |
| <i>Nave sin Capacidad de contenedores</i> | La capacidad de una nave es cargada manualmente en los datos maestros de logística. Cuando una nave no tiene su capacidad es debido a un error humano o por un tema temporal. | Se creó un nuevo archivo, el cual contiene una capacidad aproximada por Ruta marítima y naviera. En caso de que una nave no tenga su capacidad cargada, se utilizara el dato que aparece en este archivo como aproximación. |
| <i>Nave sin Flete marítimo a Puerto de Destino</i> | El Flete marítimo depende del puerto de destino, tipo de contenedor y la naviera. Cuando no se encuentra el Flete, puede ser que una de estas llaves no está creada en los datos maestros o que esa llave no se puede utilizar. | Se decidió con el área de Chartering que lo menos perjudicial es crear un flete falso de 9999 dólares y emitir una advertencia. |
| <i>Nave sin Transit Time a puerto de destino.</i> | El tiempo de tránsito depende del puerto de destino, de la naviera y de la ruta marítima. Cuando no se encuentra el tiempo de tránsito es porque la llave no está creada en los datos maestros. | Se utiliza el máximo tiempo de tránsito de las otras navieras al mismo puerto de destino. Si ninguna otra naviera va a tal puerto, entonces no se considera ninguna entrega a tal puerto. |
| <i>Nave sin entrega</i> | Por construcción, si una nave no se asocia con ninguna otra entrega, denota un problema en los datos. Puede ser que no tenga cargada los puertos de destino para esa ruta o su ETA no esté correcta. | Se mantendrán las entregas de esa en SAP como opción para el optimizador, pero para el resto de las entregas no se considerará la nave. Se emitirá una advertencia sobre estas naves. |

Tabla N°4: Resumen depuración de datos.
Fuente: Elaboración Propia.

IX. PREDICCIÓN PLANTA PRODUCTIVA

IX.1 EL PROCESO KDD

KDD (Knowledge Discovery in Database) es una metodología estándar para encontrar información en los datos y con ello generar conocimiento. Se define como un proceso no trivial de extracción de información a través de los datos, lo cual se encuentra presente de forma implícita previamente desconocida y potencialmente útil para el usuario o para el negocio.

El KDD es un proceso que consta de etapas consecutivas, y funciona de forma iterativa e interactiva. Iterativa, ya que es posible volver a cualquier etapa del proceso para ajustar parámetros o supuestos previos, e interactiva pues es el usuario experto del negocio quien tiene que estar presente para aportar con su conocimiento en la preparación de los datos y en la validación de los resultados que se obtengan.

Las etapas iniciales del proceso KDD son muy importantes. Si la preparación de los datos no está bien realizada, los resultados obtenidos en los análisis no serán confiables. Por lo tanto, hay que asegurar que se esté trabajando en un repositorio bien desarrollado, y es por esta razón que la mayor parte de los esfuerzos se emplean en la etapa de selección y preparación de los datos. Las Etapas del proceso, tal como se muestran en la *Imagen N°5* son:

- 1) Identificación del problema en estudio, teniendo un objetivo claro para el problema a resolver y cuáles son las preguntas que se quieren responder.
- 2) Selección e integración de los datos, para contar con un conjunto objetivo desde el cual obtener conocimiento. Se obtienen los datos desde los sistemas operacionales, los cuales pueden venir en diferentes formatos y en algunas oportunidades con errores, por lo cual es importante realizar una etapa de preprocesamiento.
- 3) Preparación de los datos, ya que en general, como se expresó en la etapa anterior, los datos provienen desde varias fuentes y en diferentes formatos. En esta etapa se escogen técnicas y estrategias para corregir errores en el conjunto de datos seleccionado, tratar datos faltantes y unificar formatos.
- 4) Transformación y almacenamiento de los datos, en esta etapa se pueden reducir o agrupar los datos en las características de interés. Se consolida la información y escoge una arquitectura acorde a las necesidades del problema que permita almacenarla, por ejemplo, un Data Mart.

- 5) Selección y aplicación de algoritmos de Data Mining, según el análisis que se quiera hacer, utilizando las técnicas adecuadas según las hipótesis planteadas. Los modelos y técnicas seleccionadas permitirán generar modelos de minería de datos, y con ello descubrir patrones en la información implícitos en los datos.
- 6) Interpretación y evaluación de los patrones encontrados, identificando los nuevos conocimientos y apoyándose en los expertos del negocio, para validar que el conocimiento obtenido es útil para generar acciones. Para interpretarlos, es necesario visualizarlos en diversas formas, validando los patrones y modelos de datos, de manera que generen propuestas de valor para el negocio.

IX.II MÉTODOS DE MINERÍA DE DATOS

Dentro de los métodos predictivos del Data Mining se encuentra la clasificación y la regresión, los cuales pueden ser supervisados y no supervisados. Se dice que un modelo es supervisado cuando la variable que se quiere explicar con los datos es conocida, o dicho de otra forma, que los datos tienen una "etiqueta", una regresión lineal es un ejemplo de este método. Se dice que un modelo es no supervisado cuando los datos no tienen una etiqueta, modelos que encuentran distribuciones de probabilidad, modelos de clustering o reglas de asociación son no supervisados.

La clasificación consiste en indicar a que conjunto pertenece un grupo de datos, de acuerdo a una clase predefinida. Se pretende identificar las características o atributos que hacen que un grupo de datos se vincule a un grupo a través de patrones en los datos.

La regresión es una función que le asigna a un grupo de datos un valor real. La regresión lineal es el modelo más simple, sin embargo, muchos de los problemas del mundo real no funcionan con proyecciones lineales. Las regresiones se pueden utilizar por ejemplo para predecir la demanda futura, utilizando las ventas o el consumo pasado.

Los modelos predictivos requieren ser "entrenados", utilizando un conjunto de datos de entrenamiento cuyo valor de la variable a predecir es conocido. Cuando el modelo se ajusta bien al conjunto de datos de entrenamiento, pero falla al momento de predecir en el conjunto de testeo, se dice que el modelo presenta sobreajuste u "Overfitting".

El objetivo es que el modelo entregue resultados en base a un aprendizaje, en otras palabras, que se vaya ajustando a la realidad conocida. A este tipo de modelos se les conoce como modelos de aprendizaje supervisado, debido a que los valores estimados son comparados con los resultados conocidos, y por lo tanto tienen una medida del éxito o falla de la predicción.

A continuación, se presentan los modelos y algoritmos de minería de datos utilizados en esta memoria, para la predicción de la planta productiva de un pedido. Estos algoritmos son genéricos, existen variantes de cada uno, ya sea con distintos hyper-parámetros, adaptando o combinando modelos dependiendo del tipo de problema estudiado. Cabe destacar que todos estos modelos corresponden a la categoría de aprendizaje supervisado, y si bien la predicción de la planta productiva le corresponde un modelo de clasificación, también existe una versión para problemas de regresión.

Árboles de decisión

Corresponde a uno de los métodos inductivos de aprendizaje supervisado, el cual realiza divisiones sucesivas del conjunto de datos, utilizando algún criterio de selección, manteniendo organizada su estructura de forma jerárquica, con el fin de maximizar la distancia entre los grupos de datos generados en cada iteración.

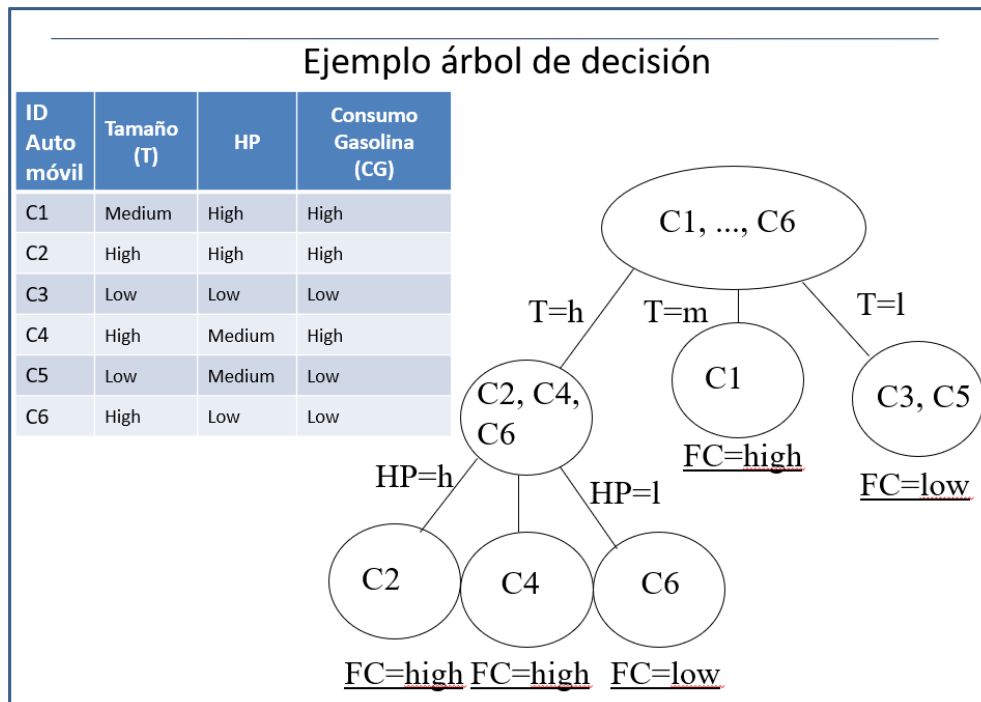


Imagen N°8: Árbol de decisión.

Fuente: Curso introduction to Data Mining – Primavera 2018.

Son una manera de representar una serie de reglas que llevan hacia una clase o valor de los datos, y se utilizan para examinarlos y realizar predicciones. Los árboles de decisión poseen una estructura formada por **Nodos**, que corresponden a los atributos o columnas que caracterizan al conjunto de datos, **Ramas**, que representan a las variables de decisión o las condiciones que cumplen los objetos para separarse unos de otros, y **Hojas** que son finalmente los conjuntos o grupos de datos resultantes de la división que realiza el algoritmo.

El algoritmo hace una clasificación discreta de los atributos, determinando a qué clase pertenece, mediante la decisión de qué rama escoger. Se asume que los grupos o clases que se formarán serán disjuntos, es decir, una instancia o grupo de datos no puede pertenecer a dos clases a la vez. Esta misma condición se cumplirá para cada partición o subárbol que se forme.

Una de las ventajas de utilizar árboles de decisión es que funcionan muy bien con variables categóricas, evitando realizar transformaciones adicionales en los datos. También permiten interpretar de buena manera las decisiones tomadas por el modelo para sus predicciones, identificando la importancia de cada atributo a partir de cientos otras variables, cosa que en algoritmos como las redes neuronales no es posible deducir. Es un método no paramétrico, es decir, no hay supuestos acerca del espacio de distribución y la estructura del clasificador.

Dentro de las desventajas de los árboles de decisión es que son muy sensible ante cambios en los datos, un pequeño cambio puede modificar ampliamente la estructura del árbol. Por lo tanto, la interpretación no es tan directa como parece. Puede presentar pérdida de información al categorizar variables continuas y a menudo tienen sobreajuste en los datos de entrenamiento.

Random Forest

El modelo de Random Forest nace para contrarrestar los problemas de sobreajuste de los árboles de decisión. El método genera múltiples árboles, cada árbol provee de una clasificación para un set de datos, y el resultado es la clase con mayor número de votos en todo el bosque (Forest). Lo anterior es conocido como "Bagging", técnica que se utiliza para reducir la varianza de las predicciones a través de la combinación de los resultados de varios trabajadores, cada uno de ellos modeladores con diferentes subconjuntos tomados de la misma población.

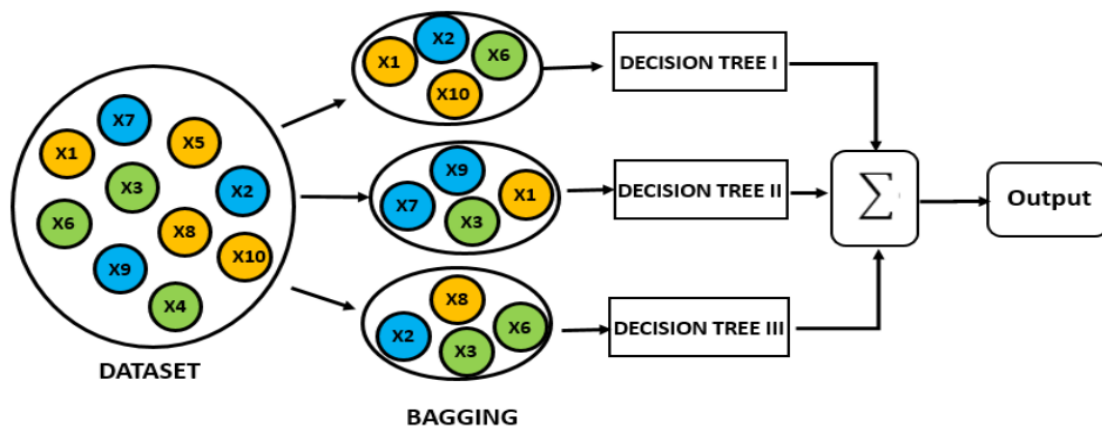


Imagen N°9: Técnica de Bagging.

Fuente: <https://bookdown.org/content/2031/> - Johanna Orellana Alvear.

Los hyper-parámetros más importante del modelo son el número de árboles, el mínimo número de muestras dentro de los nodos terminales y número máximo de nodos terminales. Aumentando el número de árboles se estabiliza el error o sobreajuste del modelo, pero usar demasiados árboles puede ser innecesariamente ineficiente computacionalmente.

Dentro de las ventajas del Random Forest es que existen muy pocas supuestas y por lo tanto la preparación de los datos es mínima, puede manejar miles de variables como entrada e identificar las más significativas. Notar que la importancia de un atributo es el incremento del error del modelo de predicción luego de que el valor de dicho atributo ha sido permutado por otro. Por su construcción, provee uno de los métodos más robustos dentro del mundo del Machine Learning, evitando el sobreajuste en los datos de entrenamiento.

Tiene la desventaja de que pierde interpretabilidad, ya que no se tiene el control de lo que hace el modelo. Debido a esto recibe el nombre de caja negra dentro de los modeladores estadísticos. Cuando los atributos están correlacionados, la medida de importancia de atributos puede estar sesgada, pues se "divide" la importancia de ambos atributos.

Redes Neuronales Artificiales

Las redes neuronales son sistemas de procesamiento de datos, cuya estructura y diseño se basa en el proceso natural del funcionamiento del cerebro. Son muy interesantes de estudiar, ya que permiten modelar eficientemente problemas complejos, en los cuales se cuenta con muchas variables o atributos predictores.

La estructura de una red neuronal está compuesta por un set de nodos interconectados, transmitiendo "señales" a través de sus conexiones, y cada una de esas conexiones o enlaces tiene un peso asociado y cada nodo una función matemática asociada.

La red neuronal funciona "aprendiendo de ejemplos", por lo que necesita un conjunto de datos iniciales, presentados en una capa de entrada, donde cada nodo de esa capa corresponde a una variable predictora. Estos nodos iniciales se conectan con otros nodos en una capa oculta, donde se les aplica una "función de activación" utilizando los pesos que tienen asociadas las conexiones entre los nodos.

Los resultados obtenidos mediante la función de activación se trasladan de un nodo a otro hacia una capa de salida, que consiste en una o más variables de respuesta del modelo. Los pesos en la red son parámetros desconocidos del modelo, y se estiman a través del aprendizaje y métodos de entrenamiento. El valor de los nodos de salida se calcula en base a los nodos de entrada y pesos iniciales de la red, combinados en la capa oculta de la red. Luego se calcula el error según el valor obtenido y el valor esperado, luego se utiliza el error para asociarlo a los nodos de la capa oculta, proporcionalmente a sus pesos. Este proceso permite ir ajustando los pesos de las conexiones en la red con el fin de reducir el error.

La imagen 10 muestra un ejemplo de Red Neuronal, con 6 nodos y 9 conexiones.

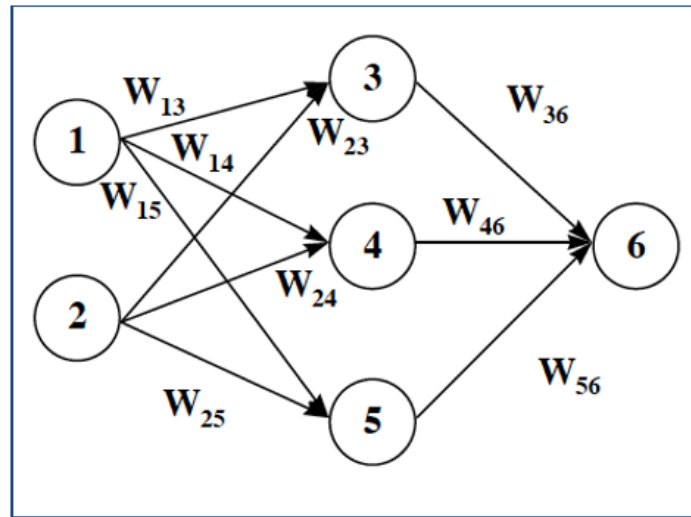


Imagen N°10: Visualización de una Red Neuronal.

Fuente: "Introduction to Data Mining and Knowledge Discovery in database"

La ventaja más relevante de las redes neuronales es su buen funcionamiento predictivo, ya que tienen una alta tolerancia a los datos anómalos y capturan muy bien la relación entre las variables y el resultado entregado. Sin embargo, a diferencia de modelos de clasificación como los árboles de decisión, es difícil interpretar las decisiones tomadas por el modelo para llegar a sus resultados, ya que funciona como una "caja negra", al igual que el Random Forest.

Support Vector Machine (SVM)

Su objetivo es encontrar la mejor función para clasificar un conjunto de datos, encontrando los hiperplanos que mejor dividan la muestra, maximizando el grado de separación entre las clases generadas.

La ventaja de utilizar este método con respecto a las redes neuronales es que resuelve el problema que tienen estas últimas de llegar a soluciones locales, ya que SVM siempre encuentra una solución global luego de haber sido entrenado, y es menos propenso a sobre ajustarse a los datos. Además, funciona bien cuando se cuenta pocos datos, no así las Redes Neuronales, las cuales solo funcionan correctamente con una cantidad importante de data.

SVM se basa en transformar los datos de entrada, desde un espacio de baja dimensión hacia uno dimensionalmente mayor. Lo anterior se realiza a partir de la elección de una función de kernel (polinomial), que se aplica sobre los datos, buscando los parámetros del modelo a través de programación cuadrática. La medida utilizada para encontrar la mejor función es maximizar el margen o la distancia entre los objetos de dos clases.

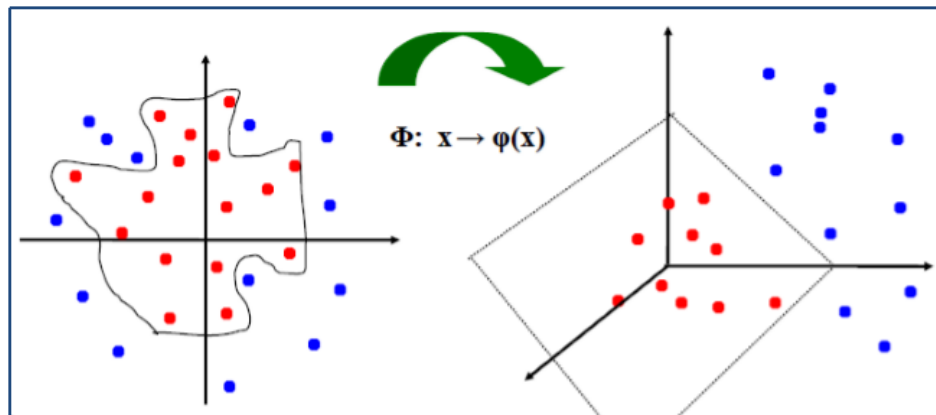


Imagen N°11: Transformación del espacio dimensional.
Fuente: Curso "Introduction to Data Mining" – Primavera 2018.

Regresión Logística

La regresión logística es similar a un modelo de regresión lineal, pero está adaptado para modelos en los que la variable dependiente es dicotómica. Los coeficientes de regresión logística pueden utilizarse para estimar la odds ratio de cada variable independiente del modelo.

La regresión logística, en concreto utiliza la distribución logística, la cual se muestra en la *Imagen N°12*. Sin embargo, la solución puede ser más estable si los predictores tienen una distribución normal multivariante.

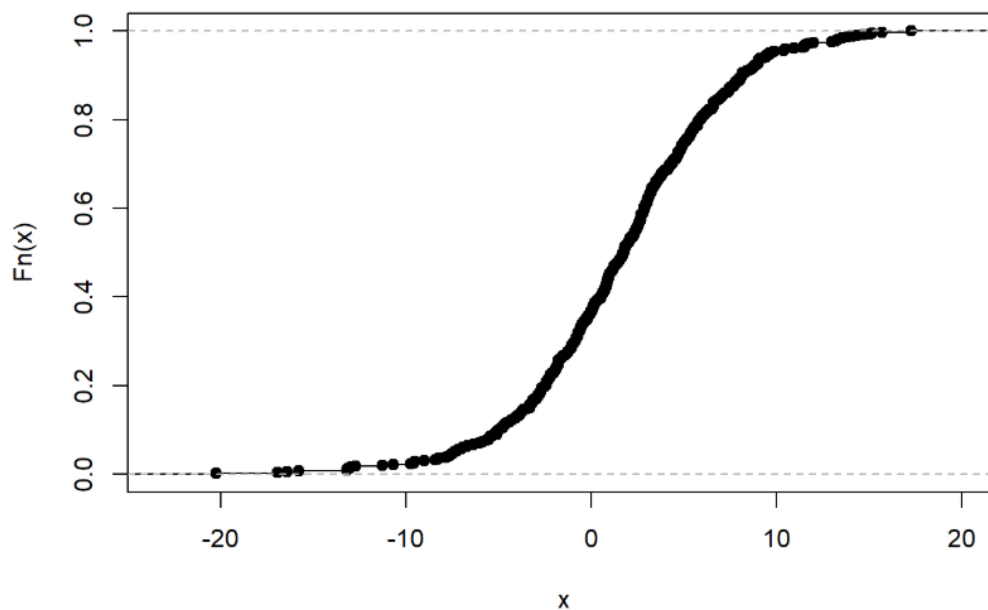


Imagen N°12: Función Distribución Logística.
Fuente: Curso "Introduction to Data Mining" – Primavera 2018.

Las ventajas de este modelo incluyen su sencilla interpretación de los coeficientes, y al ser un modelo multinivel con función de probabilidad conocida, se pueden obtener probabilidades de pertenencia a un grupo o clase para cada individuo. Además, al ser la variable dependiente dicotómica, no se necesita que se cumpla el supuesto de normalidad de los grandes números.

Dentro de las desventajas, es que se necesita una muestra grande de datos, ya que tiene como método de estimación la máxima verosimilitud, adicionalmente, al igual que con otras formas de regresión, la multicolinealidad entre los predictores puede llevar a estimaciones sesgadas y a errores estándar inflados.

Algoritmo K-NN

El método K-NN (K nearest neighbors Fix y Hodges, 1951) es un método de clasificación no paramétrico, que estima el valor de la función de densidad de probabilidad o directamente la probabilidad a posteriori de que un individuo o set de datos pertenezca a una clase u otra. En el proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables independientes o predictoras.

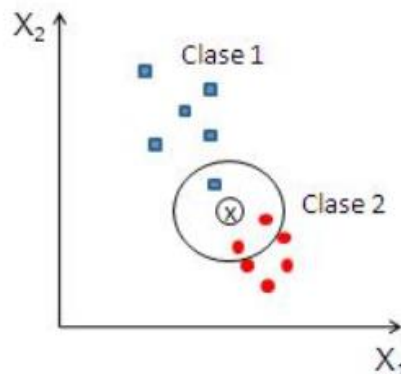


Imagen N°13: Ejemplo K-Vecinos (K=3).

Fuente: Curso "Introduction to Data Mining" – Primavera 2018.

En la *Imagen N°13* se ilustra el funcionamiento de este método de clasificación. En la figura se muestran dos clases, una formada por círculos rojos y la otra por círculos azules. En este ejemplo se han seleccionado tres vecinos, es decir, K igual a 3. De los tres vecinos más cercanos, uno de ellos pertenece a la clase 1 y dos a la clase 2. Por lo tanto, la regla 3-NN asignará la muestra a la clase 2. Generalmente se usa la distancia euclidiana para elegir los K-vecinos más cercanos.

El método KNN supone que los vecinos más cercanos nos dan la mejor clasificación y esto se hace utilizando todos los atributos; el problema de dicha suposición es que es posible que se tengan muchos atributos irrelevantes que dominen sobre la clasificación, de manera que los atributos relevantes perderían peso entre otros veinte irrelevantes.

La mejor elección de k depende fundamentalmente de los datos; generalmente, valores grandes de k reducen el efecto de ruido en la clasificación, pero crean límites entre clases parecidas. Un buen k puede ser seleccionado mediante un procedimiento de optimización. El caso especial en que la clase es predicha para ser la clase más cercana al ejemplo de entrenamiento (cuando $k=1$) es llamada *Nearest Neighbor Algorithm*, Algoritmo del vecino más cercano.

IX.1 RESULTADOS PREDICCIÓN PLANTA PRODUCTIVA

En este capítulo, se delinearán las consideraciones muestrales para formar los modelos y sus resultados, acompañados de la interpretación de estos en términos de su calidad de predicción y concordancia con la realidad. Se pondrá énfasis en el modelo de redes neuronales, debido a que fue el modelo seleccionado para ser implementado.

La selección de atributos relevantes es el proceso más importante en minería de datos, el cual consta de seleccionar el subconjunto más pequeño de atributos tal que no afecte el resultado de la clasificación y que la distribución resultante sea la más parecida a la original. Con una buena selección de atributos se puede disminuir el tiempo de procesamiento del algoritmo y mejorar la calidad de la solución disminuyendo la información irrelevante y ruidosa.

Entre los atributos seleccionados se encuentra el tipo del producto, como por ejemplo las Molduras, Celulosa UKP, Ramificados, etc., el volumen del producto, el cliente que solicitó el producto, el país de destino del producto, el tipo de material y el puerto de embarque.

El primer paso ahora que se dispone de las variables seleccionadas es diseñar los experimentos y metodologías de entrenamiento. Para poder realizar un entrenamiento eficiente hay que tener cuidado con lo altamente balanceadas que se encuentra la base de datos creada.

Modelos de machine learning como Random Forest, no lidian muy bien con set de datos de entrenamiento desbalanceados, ya que son sensibles a las proporciones de las diferentes clases. Como consecuencia, estos algoritmos tienden a favorecer a la clase con la mayor proporción de observaciones, conocida como la clase mayoritaria, lo cual puede llevar a métricas de exactitud sesgadas.

Tener un set de datos desbalanceado puede llegar a ser particularmente problemático cuando se está interesado en la clasificación correcta de una clase minoritaria, pero se llegan a altos valores de exactitud que son productos realmente de la clasificación correcta de la clase mayoritaria.

En este proyecto se utilizó la técnica del sobre muestreo que consiste en replicar la cantidad de datos de la clase más pequeña en el conjunto de entrenamiento, hasta alcanzar una proporción de 50-50. También es posible aplicar otra proporción, pero en los

modelos de Arboles de Decisión, Random Forest y Redes Neuronales funcionan mejor cuando las clases del set de entrenamiento están completamente balanceadas.

En el grafico N°6 se muestra el resultado al predecir en que planta productiva se fabricará una entrega usando los seis modelos planteados anteriormente. Para decidir cuál es el mejor modelo, se observó el Accuracy del set de testeo y el nivel de Overfitting o diferencia entre el Accuracy de los conjuntos de entrenamiento y testeo. Bajo estos criterios, el modelo de Redes Neuronales fue el de mejor desempeño.

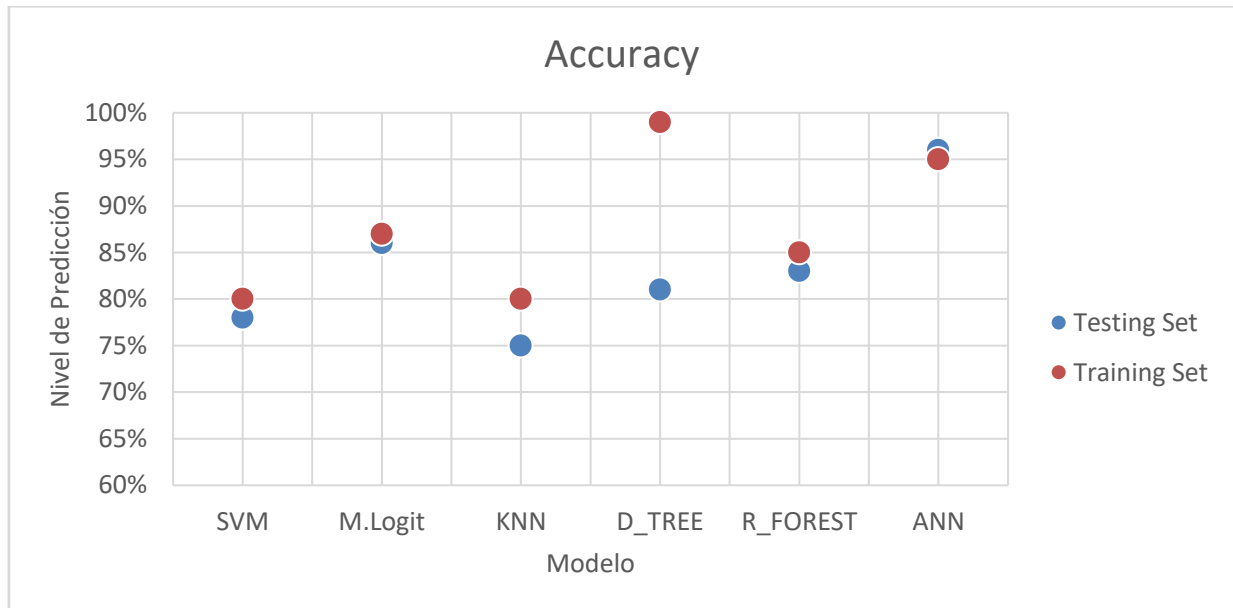


Gráfico N°6: Nivel de predicción obtenido en Conjunto de testeo y entrenamiento.
Fuente: Elaboración Propia.

Cabe destacar que, en el marco de este trabajo, no importa la interpretabilidad del modelo, es decir, no es de interés comprender por qué el modelo llegó a tal conclusión.

En la tabla N°5 se muestran los hiper parámetros utilizados en la construcción de la red. En la literatura, hasta el momento, no existe algoritmos o reglas para la selección de los hiper parámetros, hallar una buena red solo es una combinación de experiencia y suerte.

| Capa | Neuronas | Dropout | Función de activación |
|-----------|----------|---------|-----------------------|
| Input | 236 | - | - |
| 1° Oculta | 300 | 0% | Sigmoid |
| 2° Oculta | 300 | 0% | Tanh |
| 3° Oculta | 200 | 0% | Tanh |
| 4° Oculta | 100 | 0% | Tanh |
| Output | 16 | - | - |

| Epocas | BatchSide |
|--------|-----------|
| 220 | 256 |

| | |
|-----------------------|-----|
| Training set | 70% |
| Validation set | 10% |
| Testing set | 20% |

Tabla N°5: Hiper parámetros de la Red Neuronal, **Fuente:** Elaboración Propia.

X. MODELO DE OPTIMIZACIÓN

El Problema de la Mochila (conocido también como Knapsack Problem o simplemente KP) es un problema clásico de la Investigación de Operaciones y en particular de la Programación Entera. Consiste en un excursionista que debe preparar su mochila, la cual tiene una capacidad limitada y por tanto no le permite llevar todos los artículos que quisiera tener en la excursión. Cada artículo que el excursionista puede incluir en la mochila le reporta una determinada utilidad. Luego el problema consiste en seleccionar un subconjunto de objetos de forma tal que se maximice la utilidad que el excursionista obtiene, pero sin sobrepasar la capacidad de acarrear objetos. Podemos hacer la analogía con el problema del excursionista identificando a la mochila como una nave y a una entrega como un artículo donde su peso es la cantidad de contenedores.

Operativamente, el problema de asignación de entrega a buques puede ser infactible, por ejemplo, cuando la capacidad de las naves no es suficiente para embarcar todas las entregas o no hay naves con el mismo destino de una entrega. Por ello las restricciones no se modelaron de forma estrictas, sino que tienen una variable de holgura que evita la infactibilidad. El objetivo del modelo es minimizar el costo logístico y maximizar el nivel de servicio, es sencillo poner un peso con respecto al flete a incumplir con el nivel de servicio meta de Arauco, también es sencillo poner un peso con respecto al flete para sobrepasar la capacidad de una nave o cambiar de nave una entrega. Pero para los tomadores de decisión es muy difícil parametrizar el nivel de servicio con respecto a sobrepasar la capacidad de una nave o la capacidad de bodega de un puerto.

Para lograr que la parametrización del modelo sea una experiencia intuitiva para el área de Chartering, se decidió separar el modelo en dos partes, la primera minimiza el costo total, y en la segunda se maximiza el nivel de servicio al cliente.

El 70% de las restricciones se encuentran de forma implícita en el modelo a través de un conjunto de pares lógicos, el cual asocia una entrega con múltiples naves si y solo si no se rompe ninguna restricción operacional. Luego, el modelo de optimización solo podrá asignar una entrega a una de estas naves previamente asociadas. Nos referiremos a este conjunto lógico como el conjunto de pares Entrega – Nave o PAR.

A continuación, se presenta el modelo lineal, la cicla FPAN significa “Fecha Preferente de Embarque”, embarcar una entrega según este criterio significa que se debe asignar una nave cuya fecha de llegada a puerto chileno sea lo más cercano a la FPAN. La cicla FRC significa “Fill Rate Cliente” o nivel de servicio al cliente, embarcar una entrega según este criterio implica hacer llegar la entrega al lugar señalado por el cliente en la fecha que la compañía se comprometió.

Conjuntos

E = Conjunto de entregas a embarcar.

N = Conjunto de naves con formato (Nombre nave, Naviera, Número viaje, Ruta).

PAR = Conjunto de pares Entrega – Nave factibles $(i, j) \subseteq E \times N$.

PAR_{SAP} = Conjunto de pares Entrega – Nave en SAP, $PAR_{SAP} \subseteq PAR$.

P = Conjunto de puertos de embarque.

T_{PROD} = Tipos de producto con bodega especial en puerto.

E_{FPAN} = Conjunto de entregas a embarcar segun criterio FPAN.

E_{FRC} = Conjunto de entregas a embarcar segun criterio FRC.

E_{prod} = Conjunto de entregas del tipo de producto $prod \in T_{PROD}$.

S = Semanas de embarque segun ETA de las naves disponibles.

S' = Semanas actual y las tres posteriores.

G = Grupo de entregas.

E_g = Conjunto de entregas que pertenecen al grupo $g \in G$.

N_s^p = Conjunto de naves que embarcan en semana $s \in S$ por el puerto $p \in P$.

Parámetros del modelo

Cap_Nave_j = Capacidad de la nave $j \in N$ segun contrato naviero.

$Holgura_Cap_Nave_j$ = Capacidad extra de la nave $j \in N$ facilmente gestionable.

$Cap_Pto_p^{prod}$ = Capacidad contenedores semanales de puerto de embarque $p \in P$,

para el tipo de producto $prod \in T_{PROD}$.

CC_i = Cantidad de contenedores de la entrega $i \in E$.

CC_i^s = Cantidad de contenedores de la entrega $i \in E$ con nave en la semana $s \in S$.

$CC_{i,j}^s$ = Cantidad de contenedores del par Entrega – Nave $(i, j) \in PAR$ de la semana $s \in S$.

$CC_OKFPAN_{i,j}^s$ = Cantidad de contenedores del par Entrega – Nave $(i, j) \in PAR$ de la semana $s \in S$ que se encuentran OK segun criterio FPAN.

$CC_OKFRC_{i,j}^s$ = Cantidad de contenedores del par Entrega – Nave $(i, j) \in PAR$ de la semana $s \in S$ que se encuentran OK segun criterio FRC.

FPAN_META = Nivel de servicio meta de Arauco segun FPAN.

FRC_META = Nivel de servicio meta de Arauco segun FRC.

α = Flexibilidad del flete maritimo en maximizacion del Fill Rate .

Costo_Logistico* = Costo logístico óptimo del modelo de minimización de flete.

Costo_Cap_Pto* = Costo optimo de sobrepasar capacidad bodegas portuarias según modelo de minimización de flete.

CostoCapNave_Leve* = Costo óptimo de usar la holgura de una nave.

CostoCapNave_Rebasada* = Costo óptimo de sobrepasar la capacidad de una nav

Costo_Cambio_SAP* = Costo óptimo de no mantener asignación en SAP .

Parámetros de costo

$Flete_i^j$ = Costo maritimo de embarcar entrega $i \in E$ en el nave $j \in N$.

$Costo_Portuario_i^j$ = Costo portuario de embarcar entrega $i \in E$ en el nave $j \in N$.

$Costo_Terrestre_i^j$ = Costo terrestre de embarcar entrega $i \in E$ en el nave $j \in N$.

$Costo_dias_FPAN_i^j$ = Costo de lejanía ETA – FPAN del par Entrega – Nave $(i, j) \in PAR$.

$Costo_dias_FRC_i^j$ = Costo de lejanía ETD – FRC del par Entrega – Nave $(i, j) \in PAR$.

Costo_FPAN_META = Costo de incumplir en un CC en nivel de FPAN meta

Costo_FRC_META = Costo de incumplir en un CC en nivel de FRC meta.

$Costo_Cap_Pto_p^{prod}$ = Costo de sobrepasar capacidad bodega del puerto de embarque $p \in P$, del producto $prod \in T_PROD$.

CostoCapNave_Leve = Costo de pasar la capacidad de una nave levemente en un CC.

CostoCapNave_Rebasada = Costo de sobrepasar la capacidad de una nave en un CC.

Costo_Cambio_SAP_i^j = Costo cambiar asignacion en SAP de la entrega i a la nave j por CC.

Variables de decisión

$$X_i^j = \begin{cases} 1 & \text{si se asigna entrega } i \in E \text{ en nave } j \in N. \\ 0 & \text{si no} \end{cases}$$

$$Y_Cambio_Nave_SAP_i^j = \begin{cases} 1 & \text{si cambia asignación de entrega } i \in E \text{ en nave } j \in N \text{ de SAP.} \\ 0 & \text{si no} \end{cases}$$

Y_Cap_Pto_{p,s}^{prod} = Contenedores sobrepasados en puerto p ∈ P, en la semana s

∈ S, para el tipo de producto prod ∈ T_PROD.

YCapNave_Leve_j = Cantidad de contenedores pasados en nave j ∈ N.

YCapNave_Rebasada_j = Cantidad de contenedores sobrepasados en nave j ∈ N.

Y_FPAN_s = Cantidad de CC incumplidos segun Fill Rate meta FPAN en semana s ∈ S'.

Y_FRC = Cantidad de CC incumplidos segun meta FRC.

X.1 MODELO DE MINIMIZACIÓN DE FLETE

1. Asignar todas las entregas en una única nave.

$$\sum_{j \in N} X_i^j = 1 \quad \forall i \in E$$

2. Respetar la capacidad de las naves.

$$\sum_{i \in E} CC_i * X_i^j \leq Cap_Nave_j + YCapNave_Leve_j + YCapNave_Rebasada_j \quad \forall j \in N$$

3. Definición de sobrepasar levemente las naves.

$$YCapNave_Leve_j \leq Holgura_Cap_Nave_j \quad \forall j \in N$$

4. No sobrepasar la capacidad de bodegas especiales de los puertos de embarque.

$$\sum_{i \in E_{prod}} \sum_{j \in N_s^p} CC_i * X_i^j \leq Cap_Pto_p^{prod} + Y_Cap_Pto_{p,s}^{prod} \quad \forall p \in P, s \in S, prod \in T_PROD$$

5. Mantener asignación en SAP.

$$X_i^j = 1 - Y_Cambio_Nave_SAP_i^j \quad \forall (i, j) \in PAR_SAP$$

6. Naturaleza de las variables.

$$X_i^j \in \{0,1\} \quad \forall (i,j) \in \text{PAR}$$

$$Y_{\text{Cambio_Nave_SAP}}^j \in \{0,1\} \quad \forall (i,j) \in \text{PAR}$$

$$Y_{\text{Cap_Pto}}^{\text{prod}} \geq 0 \quad \forall p \in P, \text{prod} \in T_{\text{PROD}}, s \in S$$

$$Y_{\text{CapNave_Leve}}_j \geq 0 \quad \forall j \in N$$

$$Y_{\text{CapNave_Rebasada}}_j \geq 0 \quad \forall j \in N$$

Función objetivo

$$\begin{aligned} \min \sum_{i \in E} \sum_{j \in N} X_i^j * & \left(\text{Flete}_i^j + \text{Costo}_{\text{Portuario}}^j + \text{Costo}_{\text{Terrestre}}^j \right) \\ & + \sum_{i \in E} \sum_{j \in N} \sum_{s \in S} Y_{\text{Cambio_Nave_SAP}}^j * \text{Costo_Cambio_SAP}_i^j * C_{i,j}^s \\ & + \sum_{p \in P} \sum_{s \in S} \sum_{\text{prod} \in T_{\text{PROD}}} Y_{\text{Cap_Pto}}^{\text{prod}} * \text{Costo_Cap_Pto}_p^{\text{prod}} \\ & + \sum_{j \in N} Y_{\text{CapNave_Leve}}_j * \text{CostoCapNave_Leve} \\ & + \sum_{j \in N} Y_{\text{CapNave_Rebasada}}_j * \text{CostoCapNave_Rebasada} \end{aligned}$$

Los costos de Flete, portuarios y terrestres corresponden a costos reales de la compañía. Como las tarifas portuarias y terrestres están en pesos chilenos y el Flete está en dólares, se transforma todo a dólares usando la tasa de cambio del banco central. Notar que, debido a lo anterior, un cambio en la tasa de cambio del peso chileno a dólar puede afectar la solución del optimizador, dándole prioridad a embarcar en naves cuyo puerto de zarpe sea el más barato o el más cercano a la planta productiva de la entrega.

El costo de cambiar una entrega de nave también es un costo real, pero estimado bajo algunos supuestos, el cual representa el costo de los movimientos de contenedores que implica el cambio de nave. Por ejemplo, si una entrega tiene asignada una nave con Lirquén como puerto de embarque, entonces para cambiar la entrega a otra nave cuyo puerto de embarque es Coronel, entonces se debe incurrir en un costo terrestre por el traslado, además, si es que la entrega ya se encuentra consolidada en un contenedor, también puede haber movimientos de grúas involucrados.

Los supuestos del costo de cambio de nave tienen que ver con la posición y estado actual de la entrega, las cuales no se saben con total exactitud al momento de hacer el cambio.

Por regla general del departamento de logística, las entregas no pueden ser despachadas a puerto antes de 21 días del zarpe de la nave, por lo tanto, se asume que, si una entrega está a más de 21 días del zarpe de su nave, no tienen ningún costo asociado al cambio de puerto.

Por otra parte, hay productos que se consolidan en un contenedor en la planta productiva y otros en el puerto, la consolidación se puede hacer con el contenedor montado en el camión, por lo tanto, no se tiene que recurrir a una grúa, solo se indemniza al conductor del camión por el tiempo que tome la consolidación, lo cual es mucho más barato. De esta forma, a través de una serie de reglas entre los días que faltan para el zarpe y estado de consolidación de la entrega se calcula el costo de cambiar una entrega de nave (lo cual puede implicar un cambio en el puerto de embarque).

El costo de sobrepasar la capacidad de contrato de las naves se divide en dos, un costo por sobrepasar levemente la nave y otro por sobrepasar por mucho la capacidad de las naves. El límite entre estas definiciones está dado por el parámetro **Holgura_Cap_Nave**, el cual refleja la cantidad extra de espacio, en cantidad de contenedores, que es fácilmente gestionable para Chatering. Esta holgura depende de la ruta marítima y de la empresa Naviera que opera la nave, embarcar más que esta holgura es muy difícil, por lo tanto, se tratara como una infactibilidad al igual que sobrepasar la capacidad de los puertos.

Aunque sea "Sencillo" conseguir un poco más de la capacidad contratada para algunas naves, no es deseable hacerlo, pues primero hay que cumplir los contratos con todas las navieras. Finalmente se definen tres niveles de penalidades en la función objetivo, el primer nivel corresponde a los costos reales (Flete marítimo, costo portuario, terrestre y cambio de nave), el nivel 2 corresponde a situaciones indeseadas (Costo por sobrepasar levemente las naves) y el nivel 3 son penaltis que evitan infactibilidades (Costo de sobrepasar capacidad de las naves y de los puertos).

Nivel 1 << Nivel 2 << Nivel 3

Queda a criterio de los tomadores de decisión que hacer cuando se incurre en una infactibilidad y cómo gestionar los penaltis dentro del nivel 3, pues el escenario más deseable o gestionable dependerá de la fecha en que se tome la decisión.

X.II MODELO MAXIMIZACIÓN NIVEL DE SERVICIO

X.II Asignar todas las entregas en una única nave.

$$\sum_{j \in N} X_i^j = 1 \quad \forall i \in E$$

XI.II Respetar la capacidad de las naves.

$$\sum_{i \in E} CC_i * X_i^j \leq Cap_Nave_j + YCapNave_Leve_j + YCapNave_Rebasada_j \quad \forall j \in N$$

XII.II Definición de sobrepasar levemente las naves.

$$YCapNave_Leve_j \leq Holgura_Cap_Nave_j \quad \forall j \in N$$

XIII.II No sobrepasar la capacidad de bodegas especiales de los puertos de embarque.

$$\sum_{i \in E_{prod}} \sum_{j \in N_s^p} CC_i * X_i^j \leq Cap_Pto_p^{prod} + Y_Cap_Pto_{p,s}^{prod} \quad \forall p \in P, s \in S, prod \in T_PROD$$

XIV.II Mantener asignación en SAP.

$$X_i^j = 1 - Y_Cambio_Nave_SAP_i^j \quad \forall (i, j) \in PAR_SAP$$

XV.II Cumplir con nivel de asignación por FPAN semanal.

$$\sum_{i \in E_{FPAN}} \sum_{j \in N} CC_OKFPAN_{i,j}^s * X_i^j \geq FPAN_META * \sum_{i \in E_{FPAN}} CC_i^s - Y_FPAN_s \quad \forall s \in S'$$

XVI.II Cumplir con nivel de servicio al cliente.

$$\sum_{i \in E_{FRC}} \sum_{j \in N} \sum_{s \in S'} CC_OKFRC_{i,j}^s * X_i^j \geq FRC_META * \sum_{i \in E_{FPAN}} \sum_{s \in S'} CC_i^s - Y_FRC$$

XVII.II Coordinación con modelo de minimización de flete.

$$\sum_{i \in E} \sum_{j \in N} X_i^j * (Flete_i^j + Costo_Portuario_i^j + Costo_Terrestre_i^j) \leq \alpha Costo_Logistico^*$$

$$\sum_{i \in E} \sum_{j \in N} \sum_{s \in S} Y_Cambio_Nave_SAP_i^j * Costo_Cambio_SAP_i^j * C_{i,j}^s \leq \alpha Costo_Cambio_SAP^*$$

$$\sum_{p \in P} \sum_{s \in S} \sum_{prod \in T_PROD} Y_Cap_Pto_{p,s}^{prod} * Costo_Cap_Pto_p^{prod} \leq \alpha Costo_Cap_Pto^*$$

$$\sum_{j \in N} YCapNave_Leve_j * CostoCapNave_Leve \leq \alpha CostoCapNave_Leve^*$$

$$\sum_{j \in N} YCapNave_Rebasada_j * CostoCapNave_Rebasada \leq \alpha CostoCapNave_Rebasada^*$$

XVIII.II Naturaleza de las variables.

$$X_i^j \in \{0,1\} \quad \forall (i,j) \in \text{PAR}$$

$$Y_{\text{Cambio_Nave_SAP}}^j \in \{0,1\} \quad \forall (i,j) \in \text{PAR}$$

$$Y_{\text{Cap_Pto}}^{\text{prod}} \geq 0 \quad \forall p \in P, \text{prod} \in T_PROD, s \in S$$

$$Y_{\text{CapNave_Leve}} \geq 0 \quad \forall j \in N$$

$$Y_{\text{CapNave_Rebasada}} \geq 0 \quad \forall j \in N$$

$$Y_{\text{FPAN}}, Y_{\text{FRC}} \geq 0$$

Función objetivo

$$\begin{aligned} \min \sum_{s \in S} Y_{\text{FPAN}}^s * \text{Costo_FPAN_META} + Y_{\text{FRC}} * \text{Costo_FRC_META} \\ + \sum_{i \in E_{\text{FPAN}}} \sum_{j \in N} X_i^j * \text{Costo_dias_FPAN}_i^j + \sum_{i \in E_{\text{FRC}}} \sum_{j \in N} X_i^j * \text{Costo_dias_FRC}_i^j \end{aligned}$$

Los costos por incumplir en un contenedor el nivel de servicio por FRC y FPAN son idénticos, pues para la compañía ambos tienen la misma importancia. Respecto a los días de alejamiento, tienen más peso por día el criterio FRC, pues hay más incertidumbre en la fecha de llegada a destino (Definición Criterio FRC) que en el zarpe de la nave (Definición criterio FPAN) así que es más relevante acercarse un día más a la fecha prometida de llegada que un día en la fecha de embarque prometida.

Por lo tanto, al igual que en el modelo de minimización de costos, en la función objetivo del modelo de maximización del nivel de servicio también se encuentran tres niveles. En el primer nivel se encuentra el costo por incumplir en un contenedor el nivel de servicio (Costo FRC y FPAN meta), en el segundo nivel está el costo por días de lejanía de la fecha prometida al cliente y en el tercer nivel los días de lejanía a la fecha de zarpe prometida.

Nivel 1 >> Nivel 2 > Nivel 3

El parámetro α permite comunicar ambos modelos. Si el nivel de servicio óptimo es muy bajo, se podría mejorar aumentando el costo logístico, esto se hace con un $\alpha > 1$. A priori se pensó en un α igual para todas las restricciones de coordinación (Restricción 8 en el modelo), pero podrían ser distintas.

X.III RESTRICCIONES IMPLÍCITAS, CONJUNTO DE PARES

En el horizonte de los tomadores de decisión, se pueden encontrar aproximadamente cinco mil entregas y doscientas naves distintas, generando alrededor de un millón de variables enteras. La construcción de los pares tiene como objetivo disminuir la cantidad de variables enteras en el modelo lineal.

Cabe destacar que una nave puede ser operada por más de una naviera, es decir, cada naviera gestiona una parte del espacio total de una nave. Dependiendo de la naviera, se tienen tiempos de tránsito distintos y fletes marítimos distintos, gatillado principalmente por los distintos transbordos y rutas que realizan las empresas navieras con sus contenedores. A continuación, se describen las restricciones implícitas más importantes:

Puerto de destino

Cada entrega tiene un puerto de destino, previamente acordado con el cliente, el cual se actualiza con el área de ventas. Por otro lado, las naves tienen puertos de arribo, los cuales vienen dado por sus rutas marítimas. Entonces una entrega se puede embarcar en una nave solo si el puerto de destino de la entrega se encuentra dentro de la ruta de la nave.

Fecha Productiva

Cada entrega tiene una Fecha Preferente de Embarque o FPAN. Esta fecha es la suma de la fecha de término de producción, el tiempo que la entrega se consolida en un contenedor y los días que debe pasar por aduanas. En la operación, no se puede asignar una entrega a una nave cuya fecha de arribo o ETA sea menor a la FPAN menos dos días, estos dos días están dentro de la holgura usada al calcular la FPAN de una entrega y además considerando que una nave demora en promedio dos días en zarpar, se decidió operar bajo esta condición de dos días.

En el esquema N°1 de la siguiente hoja ejemplifica estas dos primeras restricciones. En el ejemplo aparece una entrega y ocho naves en tablas distintas, cruzar estas dos tablas es equivalente a usar un *JOIN*, como se conoce en base datos, donde la llave para cruzar ambas tablas sería el puerto de destino.

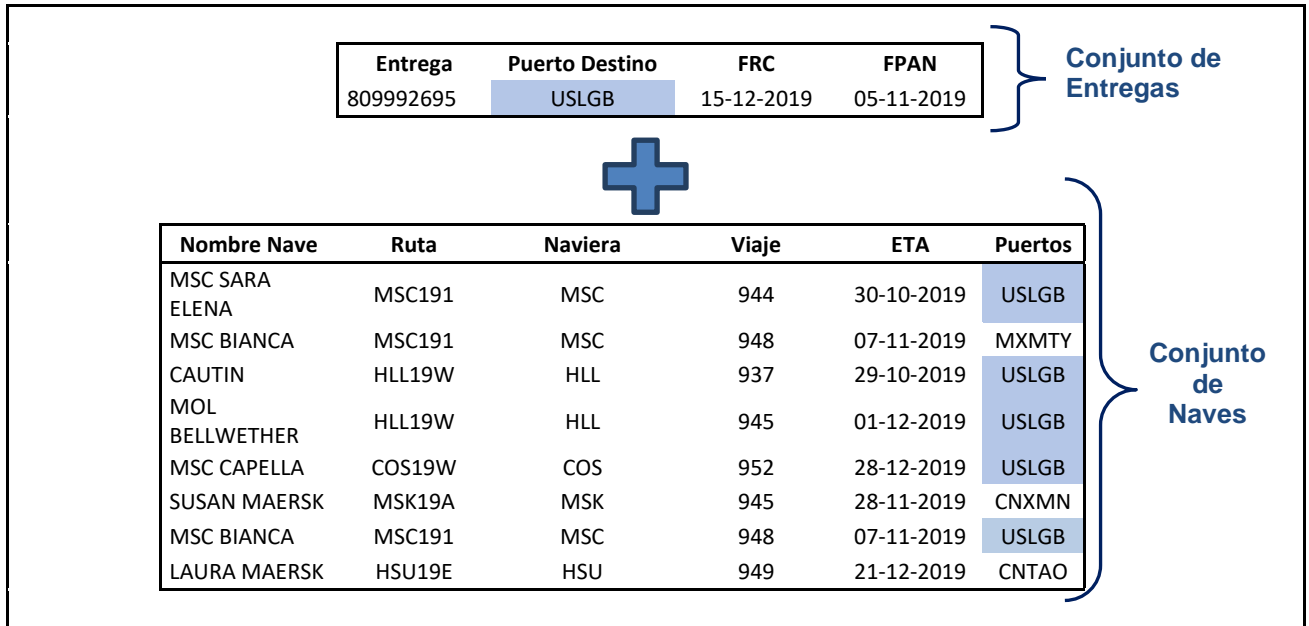
Mantener Puerto de Embarque

Existe un conjunto de entregas que ya fueron despachadas a su puerto de embarque, dado que la empresa no tiene la capacidad para gestionar el transporte de un contenedor de un puerto de embarque a otro a puertas del zarpe de la nave, estas entregas no se pueden cambiar de puerto. Entonces si una entrega ya está en algún puerto de embarque y está a menos de 7 días del zarpe de su nave respectiva, solo se asociará dicha entrega con naves que recalen en el mismo puerto.

Entregas Con Bloqueo Financiero

Cada cliente tiene una línea de crédito con Arauco, cuando se alcanza el límite de una línea de crédito, todos los pedidos de dicho cliente quedan con un estado de bloqueadas. El

estado bloqueado de una entrega que esta se fabrique y/o que se embarque. En la operación Chartering no asigna estas entregas a ninguna nave, pero, para efectos de la programación del modelo, estas entregas se asociarán con una única nave ficticia con el nombre de "No Asignada: Entrega Bloqueada".



| Entrega | Puerto Destino | FRC | FPAN | Nombre Nave | Ruta | Naviera | Viaje | ETA |
|-----------|----------------|------------|------------|----------------|--------|---------|-------|------------|
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MSC SARA ELENA | MSC191 | MSC | 944 | 30-10-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MSC BIANCA | MSC191 | MSC | 948 | 07-11-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | CAUTIN | HLL19W | HLL | 937 | 29-10-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MOL BELLWETHER | HLL19W | HLL | 945 | 01-12-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MSC CAPELLA | COS19W | COS | 952 | 28-12-2019 |



| Entrega | Puerto Destino | FRC | FPAN | Nombre Nave | Ruta | Naviera | Viaje | ETA |
|-----------|----------------|------------|------------|----------------|--------|---------|-------|------------|
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MSC BIANCA | MSC191 | MSC | 948 | 27-11-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MOL BELLWETHER | HLL19W | HLL | 945 | 21-12-2019 |
| 809992695 | USLGB | 15-12-2019 | 05-11-2019 | MSC CAPELLA | COS19W | COS | 952 | 28-12-2019 |

Esquema N°1: Elaboración del conjunto de pares Entrega-Nave. Restricción Puerto de destino y Fecha productiva.

Fuente: Elaboración Propia.


Grupo de entregas

Existen entregas que deben compartir el mismo contenedor. Estas entregas son pedidos que un mismo cliente realizo, pero por temas temporales y sistémicos se ingresaron al sistema como entregas independientes. Toda entrega que forman parte de un grupo tiene un número de grupo asociado, y todas las entregas con el mismo número de grupo deben ser embarcadas en la misma nave. Esta restricción modelada linealmente quedaría de la siguiente manera:

$$X_i^j = X_k^j \quad \forall j \in N, \quad \forall g \in G, \quad \forall i, k \in G_g$$

Para incluir esta restricción de forma implícita, el número de grupo de entrega paso a ser una entrega ficticia, la cual representa todas las entregas de tal grupo. De esta forma, se asociarán las naves con el Grupo de Entregas, y al momento de escoger una nave para ese grupo de entregas, implícitamente todas las entregas del grupo quedarán asignadas a la misma nave. El esquema N°2 expresa lo descrito anteriormente.

| Entrega | Grupo Entrega | Nave | Optimo |
|---------|---------------|----------------|--------|
| 123 | 19557 | MSC AVNI | 1 |
| 124 | 19557 | MSC AVNI | 1 |
| 125 | 19557 | MSC AVNI | 1 |
| 123 | 19557 | MOL BELLWETHER | 0 |
| 124 | 19557 | MOL BELLWETHER | 0 |
| 125 | 19557 | MOL BELLWETHER | 0 |



| Entrega | Grupo Entrega | Nave | Optimo |
|---------|---------------|----------------|--------|
| 19557 | 19557 | MSC AVNI | 1 |
| 19557 | 19557 | MOL BELLWETHER | 0 |

Esquema N°2: Elaboración del conjunto de pares Entrega-Nave. Restricción de Grupo de Entregas.

Fuente: Elaboración Propia.

Grupo de entregas bloqueado

Cuando un cliente alcanza el límite de su línea de crédito con Arauco, todos sus pedidos quedan bloqueados y por tanto no se pueden embarcar. Pero, por errores en el sistema, cuando las entregas pertenecen a un Grupo de Entregas, estas no siempre aparecen como bloqueadas. Entonces, si hay una entrega de algún Grupo de Entregas bloqueada, entonces todas las entregas del grupo se deben bloquear. Para efectos de modelación, estas entregas se asociarán con una única nave ficticia con el nombre de "No Asignada: Grupo de Entregas Bloqueado".

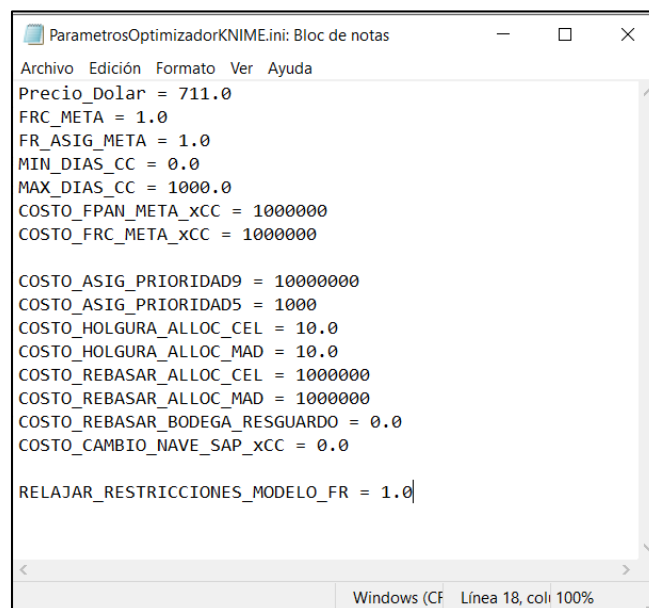
X.IV IMPLEMENTACIÓN Y MANTENCIÓN DE LA HERRAMIENTA

La herramienta se programó para ejecutarse 3 veces al día de forma automática en un servidor local, tal proceso automático comienza con la descarga de datos operacionales desde SAP y el rescate de datos que el área de Chartering maneja en sus carpetas. Una vez que todos los datos fuente estén preparados, la herramienta comienza con la depuración y transformación hasta generar la base de datos que utilizará el modelo de Redes

Neuronales, una vez ejecutado este modelo, se genera la base de datos que se utilizará en el modelo de optimización.

No es necesario ejecutar siempre el modelo de Redes Neuronales, solo basta con que se ejecute al menos una vez al mes para actualizar la Red entrenada con datos más actuales. Cuando se quiere usar la red para predecir la planta productiva de una entrega, se utilizará la última Red entrenada que se encuentre en los archivos de respaldo del sistema.

El modelo de optimización se ejecuta solo una vez por cada ejecución automática de la herramienta, pero es posible volver a ejecutar solo el modelo de optimización desde el archivo Excel donde se muestran los resultados. Desde el archivo de texto "ParametrosKnimeOptimizador.ini" se pueden cambiar los parámetros asociados a la función objetivo del modelo, con el fin de iterar entre costo y nivel de servicio sin la necesidad de ingresar al sistema.



```
Archivo Edición Formato Ver Ayuda
Precio_Dolar = 711.0
FRC_META = 1.0
FR_ASIG_META = 1.0
MIN_DIAS_CC = 0.0
MAX_DIAS_CC = 1000.0
COSTO_FPAN_META_xCC = 1000000
COSTO_FRC_META_xCC = 1000000

COSTO_ASIG_PRIORIDAD9 = 10000000
COSTO_ASIG_PRIORIDAD5 = 1000
COSTO_HOLGURA_ALLOC_CEL = 10.0
COSTO_HOLGURA_ALLOC_MAD = 10.0
COSTO_REBASAR_ALLOC_CEL = 1000000
COSTO_REBASAR_ALLOC_MAD = 1000000
COSTO_REBASAR_BODEGA_RESGUARDO = 0.0
COSTO_CAMBIO_NAVES_SAP_xCC = 0.0

RELAJAR_RESTRICCIONES_MODELO_FR = 1.0
```

Imagen N°14: Parámetros variables en una instancia de optimización.

Fuente: Elaboración propia.

En la implementación se realizaron tres clases guiadas, donde se enseñó como hacer modificaciones sencillas en el sistema, se enseñó como interactuar con el sistema usando, modificando datos de entrada como el de la imagen N°14, se enseñó, además, como el sistema utiliza los datos, como los interpreta y que supuestos usa. Finalmente se hizo entrega de un manual a nivel usuario, donde se resume todo lo visto en las clases.

El área de Chartering será la encargada de mantener actualizada la base de datos que el sistema utiliza, datos que usan regularmente en su trabajo diario, así que no implica un trabajo extra, pero si una mayor preocupación al velar por la calidad de los datos. El área de Planificación queda a cargo de hacer modificaciones más complejas en el sistema, como nuevas restricciones en el modelo o ingresar una nueva fuente de datos.

XI. RESULTADOS

XI.I AHORRO COSTO LOGÍSTICO ANUAL

Para analizar los resultados, se comparó la asignación de las entregas realizada por el equipo de Chartering con la asignación que realizó el optimizador. Las entregas que ya tienen una nave asignada se encuentran en el sistema SAP, pero hay un conjunto no menor de entregas que aún no tienen nave asignada, este conjunto de entregas representa el 10% aproximadamente del volumen total, las cuales pueden cambiar completamente la solución del optimizador.

Las entregas sin nave son aquellas que tienen algún problema grave en sus datos, como datos faltantes o inconsistencias, además de las entregas que acaban de ser creadas en el sistema. Al momento de asignar entregas, el equipo de Chartering considera las entregas sin nave, tratando de dejar espacio en las naves para hacer cabida a estas entregas. Entonces comparar solo las entregas con nave no es una comparación justa, pues el optimizador no estaría mirando las entregas sin nave.

Para hacer justa la comparación, primero se realiza una corrida del optimizador con las entregas con y sin nave, pero dejando fija la asignación de las entregas con nave, de esta forma conseguiremos el conjunto de comparación, en adelante Pool de Chartering. Luego se correrá nuevamente el optimizador con las entregas con y sin nave, permitiendo que pueda reasignar todas las entregas.

Como se verá más adelante, si se desea un mayor nivel de servicio, se requiere gastar más dinero. Para hacer los escenarios comparables, ambos deben tener el mismo nivel de servicio tanto de FPAN como de FRC, esto se logra modificando el nivel de servicio meta en los parámetros y relajando la restricción de máximo flete en el modelo de maximización de nivel de servicio. De esta forma se realizaron 6 comparaciones, usando datos de diferentes meses del año 2019.

| ID Escenario | Fecha Data Set | Costo logístico total [USD] | |
|--------------|----------------|-----------------------------|----------------|
| | | Pool de Chartering | Optimizador |
| Escenario 1 | 06-05-2019 | X1 | X1 - \$84.080 |
| Escenario 2 | 17-06-2019 | X2 | X2 - \$97.300 |
| Escenario 3 | 18-07-2019 | X3 | X3 - \$122.250 |
| Escenario 4 | 28-08-2019 | X4 | X4 - \$113.340 |
| Escenario 5 | 16-09-2019 | X5 | X5 - \$62.880 |
| Escenario 6 | 07-10-2019 | X6 | X6 - \$93.050 |

Tabla N°6: Resultado de escenarios: Ahorro en dólares.

Fuente: Elaboración propia.

La tabla 6 muestra el ahorro del costo de flete, terrestre y portuario de cada escenario, considerando un mismo nivel de servicio que el Pool de Chartering. Los resultados muestran ahorros desde \$62.880 hasta \$113.340 dólares. Para no revelar información relevante de la empresa, tanto el costo total, la cantidad de contenedores y el de nivel de servicio de cada escenario se ocultaron, dejando solo en evidencia la diferencia en costo total.

El ahorro en cada escenario no tiene un periodo de tiempo claro, así que se anualizara. En cada escenario, se dividirá el ahorro en el número de contenedores para obtener el ahorro por contenedor, luego se promediará el ahorro por contenedor de los 6 escenarios y se multiplicara por la cantidad de contenedores que se exportan en un año.

Anualizando el ahorro, se obtiene un ahorro aproximado de \$600.000 USD al año.

XI.II FLETE VS NIVEL DE SERVICIO

Para el área de logística de Arauco, siempre ha sido una interrogante de cuánto dinero se necesita gastar o se está gastando en mejorar el nivel de servicio. Esta interrogante puede ser resuelta iterando el modelo de optimización construido sobre el nivel de servicio meta. El gráfico N°7 muestra cuanto más hay que gastar para mejorar el nivel de servicio por FRC porcentualmente para un escenario o día de planificación, considerando como punto de partida el mejor nivel de servicio FRC que se puede alcanzar cuando se minimiza el costo logístico total.

A partir del gráfico, calculando la pendiente de la curva, se puede decir que se deben gastar \$12.200 USD para aumentar un 1% el nivel de servicio por FRC aproximadamente.

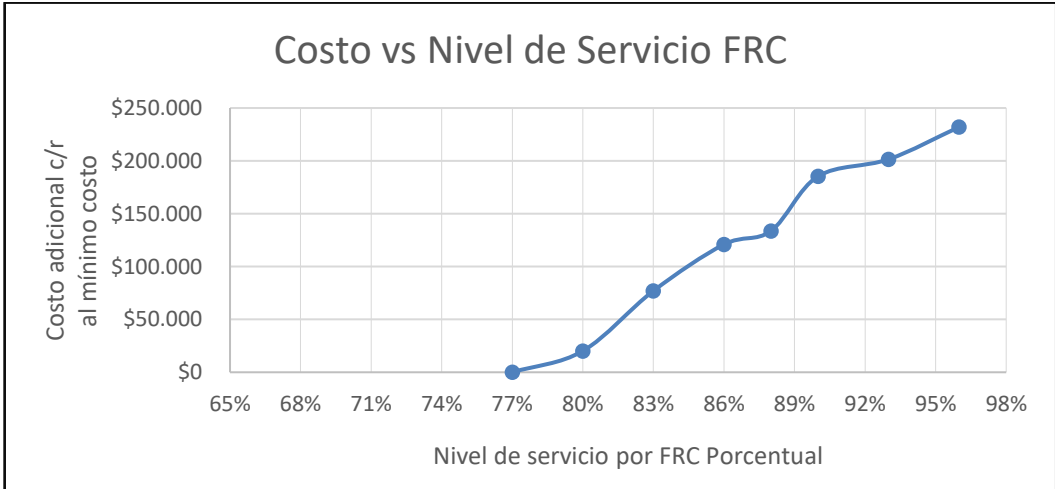


Gráfico N°7: Costo del nivel de servicio según fecha requerida por el cliente o FRC.

Fuente: Elaboración Propia.

Análogamente, para calcular el costo de mejorar una unidad porcentual el nivel de servicio por FPAN, se itera el optimizador construido variando el nivel de servicio Meta por FPAN. En el gráfico N°8 se muestran los resultados de estas iteraciones para un escenario o día de planificación, considerando como punto de partida el mejor nivel de servicio por FPAN que se puede alcanzar cuando se minimiza el costo logístico total. La pendiente estimada nos dice que se debe gastar \$6.700 USD en mejorar en un 1% el nivel de servicio por FPAN.

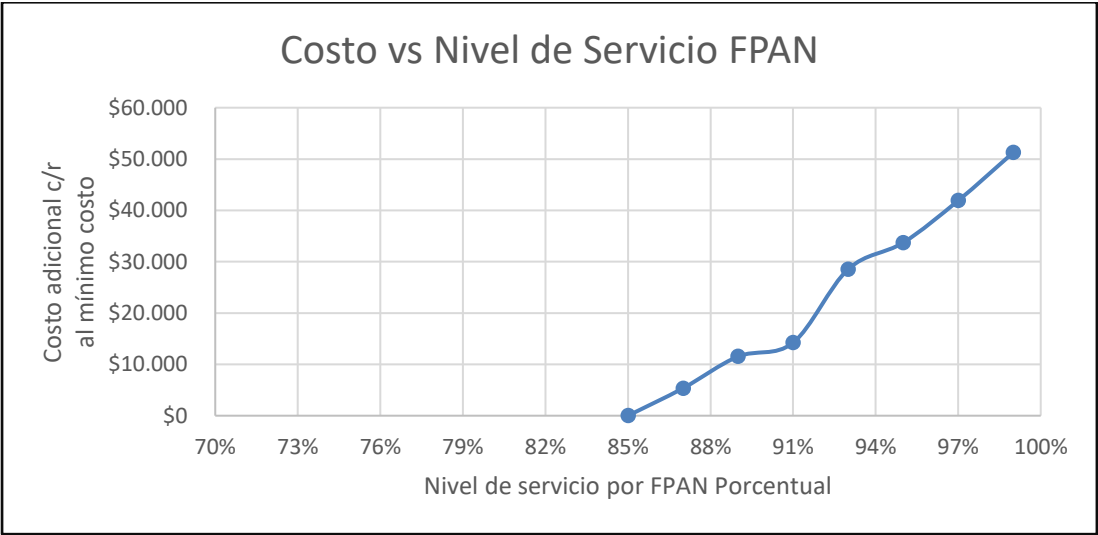


Gráfico N°8: Costo del nivel de servicio según la fecha preferente de embarque o FPAN.

Fuente: Elaboración Propia.

Los resultados que se muestran en los gráficos N°7 y N°8 no son concluyentes, pues solo consideran las entregas de un escenario, es decir, los gráficos solo muestran una parte del año. Para alcanzar resultados más certeros, es necesario realizar este mismo ejercicio en distintos instantes del año, al menos cada dos semanas, periodo en el cual cambiarán la mayoría de las entregas a ser asignadas.

Se espera que en general se deba incurrir en un mayor gasto en mejorar el nivel de servicio por FRC debido a que se está pagando para que el contenedor llegue en menos tiempo al puerto de destino, esto implica pagar por viajes directos o con solo un transbordo. En cambio, cuando se embarca por FPAN solo se paga extra por embarcar en navieras más costosas, sin importar si el tiempo en llegada a destino aumenta (en algunos casos) o disminuye muy poco, solo importa poder embarcarla en una nave que zarpe dentro del rango aceptable.

XI.III ANÁLISIS FUNCIÓN OBJETIVO DEL NIVEL DE SERVICIO

Respecto a la función objetivo del modelo de maximización del nivel de servicio, se intentó en primer lugar considerar solo las entregas asignadas a nave cuya llegada estimada este dentro del rango de cumplimiento. Pero se tenía el problema que no consideraba los días de acercamiento a la fecha requerida por el cliente o de embarque, FRC y FPAN respectivamente, provocando, por ejemplo, que fuera indistinto para el optimizador asignar la entrega a una nave cuya fecha de llegada estimada estos dos días antes de la fecha prometida o asignarla a otra nave cuya fecha estimada de llegada es ocho días después de la prometida.

Los días de diferencia entre la FRC y la fecha estimada de llegada a destino o ETD, es relevante en el sentido de la incertidumbre en el tiempo de viaje provocada por posibles marejadas, mal clima, fecha de arribo de la nave en puerto chileno, colapso de algunos puertos intermedios, arribar en menos puertos que los planificados o incluso una falla mecánica de la nave. Debido a esta incertidumbre, es deseable asignar una entrega en naves de tal forma de que la diferencia FRC – ETD sea la menor posible.

La imagen N°15 se muestra lo explicado anteriormente, en la imagen, la mejor nave para embarcar la entrega es la Nave 1, pues está dentro del rango OK FRC y la diferencia entre las fechas FRC y ETD es mínima.

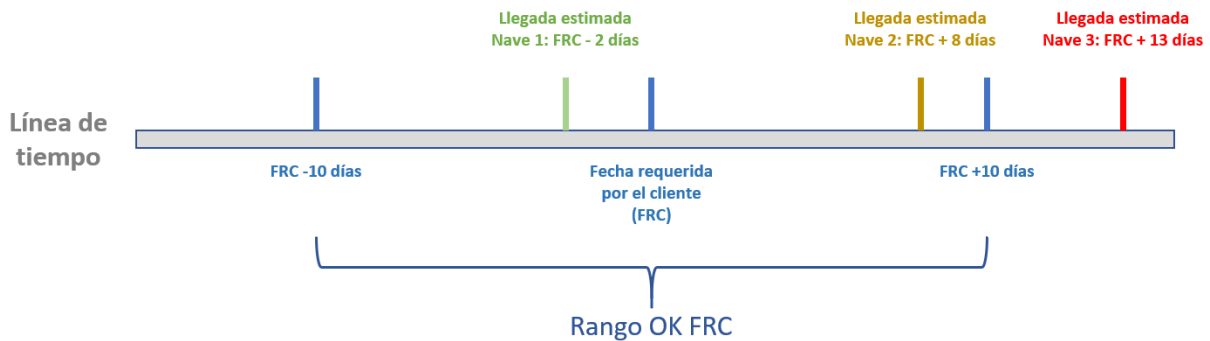


Imagen N°15: Selección de la mejor nave para el nivel de servicio.

Fuente: Elaboración propia.

Considerar la diferencia entre las fechas FRC y ETD también ayuda a que las entregas que ya están atrasadas, es decir, que es imposible que lleguen en el rango OK FRC, se embarquen lo antes posible.

Las entregas que se embarcan según el criterio de FPAN no tienen la incertidumbre del tiempo de tránsito, pues al momento de embarcar la entrega ya se cumplió con lo prometido. En este caso, la incertidumbre se encuentra en la fecha de arribo de la nave en puerto chileno, la cual también está presente en las entregas de FRC, así que acercarse a la

fecha de embarque o FPAN en un día no puede ser igual de importante que acercarse en un día a la fecha requerida por el cliente o FRC.

Tampoco es lo mismo mejorar los días de alejamiento con respecto a la fecha prometida de una entrega que tiene 2 contenedores versus otra entrega que tiene 10 contenedores, así que se debe corregir por la cantidad de contenedores. Además, se corrige por la lejanía de la entrega con respecto a la fecha actual, pues es más importante mejorar la asignación de las entregas que se deben embarcar en las próximas 2 o 3 semanas que aquellas que se pueden embarcar en la cuarta semana futura. Esto último se debe a que no se conocen todas las entregas que se van a embarcar a la cuarta semana, y también la cantidad de naves, rutas y tiempos de tránsito puede cambiar, cambiando las condiciones y el espacio factible del optimizador y haciendo probable que las entregas cuyo embarque se planifican a más de cuatro semanas sean reasignadas a otra nave.

Finalmente, se decidió construir la función objetivo del modelo de maximización de nivel de servicio de tal forma que primero intente poner a las entregas dentro del rango del nivel de servicio prometido (± 10 días para FRC y -2 a $+6$ días para FPAN), luego que mejore la cantidad de días entre FRC y ETD por unidad de contenedores, y en el último nivel, que mejore los días de diferencia entre la FPAN y el Zarpe de la nave por unidad de contenedores.

En la tabla 7 se muestra el resultado de operar con distintas funciones objetivo. Los Días por Contenedor corresponde al promedio de la diferencia entre los días de desfase con respecto a la fecha prometida multiplicado por la cantidad de contenedores de la entrega, la cual se construye de la forma descrita en los párrafos anteriores, pero sin considerar las correcciones por tipo de Embarque (FPAN o FRC) y de lejanía con respecto a la fecha actual.

Notar que minimizar los Días por Contenedor por si solos no son suficiente para garantizar que se está maximizando el nivel de servicio, tampoco es suficiente maximizar la cantidad de contenedores que caen dentro del rango de la fecha prometida, pues se puede mejorar la posibilidad de caer efectivamente dentro del rango (Recordar que se trabaja con una fecha de llegada estimada). De esta forma se, se concluyó que la combinación de los dos componentes asegura que efectivamente se está haciendo lo mejor posible para maximizar el nivel de servicio entregado al cliente.

| MAX Entregas en Rango y MIN Días por Contenedor | | MAX Entregas en Rango | | MIN Días por Contenedor | |
|-------------------------------------------------------|-------|-----------------------|-------|-------------------------|-------|
| FPAN | 92,4% | FPAN | 92,4% | FPAN | 83,4% |
| FRC | 95,9% | FRC | 95,9% | FRC | 92,2% |
| Días CC | 3,5 | Días CC | 5,4 | Días CC | 2,7 |
| Desv. Días CC | 6,2 | Desv. Días CC | 6,5 | Desv. Días CC | 5,6 |

Tabla N°7: Análisis función objetivo modelo de nivel de servicio.

Fuente: Elaboración propia.

XI.V REASIGNACIONES

Para obtener un ahorro en costo logístico o mejorar el nivel de servicio, se necesita reasignar una cierta cantidad de entregas a otras naves. Entre más reasignaciones se hagan, más beneficio se puede obtener, pero tener una alta cantidad de reasignaciones cada día no es deseable para el área Chartering. Una reasignación puede implicar trabajo de gestión y coordinación adicional con camiones, puertos y empresas navieras.

Lo anterior motiva a analizar la cantidad de reasignaciones. Para ello se itera el optimizador restringiendo el número de reasignaciones posibles agregando la siguiente restricción:

$$\sum_{(i,j) \in \text{PARES}} Y_{\text{Cambio_Nave_SAP}_i^j} \leq \text{Maximas_Reasignaciones_Permitidas}$$

Los resultados para un escenario se muestran en el gráfico 9, se observa el 98% del ahorro se consigue con la mitad de las reasignaciones totales. Esto deja en evidencia la necesidad de incluir una tercera etapa en el modelo de optimización, en donde se minimice la cantidad de reasignaciones sujeto a que se debe alcanzar un mínimo porcentaje de ahorro y de nivel de servicio.

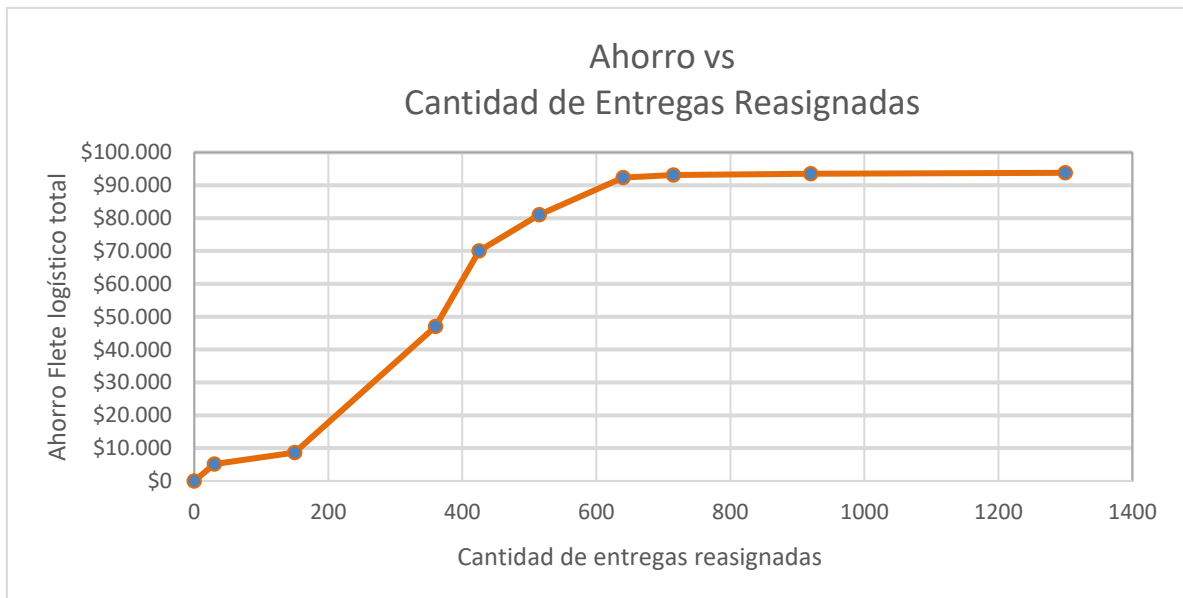


Gráfico N°9: Análisis de Pareto, ahorro total alcanzado limitando reasignación de entregas.

Fuente: Elaboración Propia.

Revisar la cantidad de reasignaciones entre Navieras y Rutas Marítimas puede ayudar a sacar conclusiones relevantes para la negociación de los nuevos contratos navieros y de las operaciones de Chartering. En la tabla 8 se muestra el flujo de entregas reasignadas por

ruta, construido a partir de una iteración del optimizador en uno de los escenarios presentados en la tabla 6.

A partir de la tabla 8, se puede observar que las asignaciones de entregas a naves con ruta 1, 2 y 16 están muy cercanas al óptimo o que en las rutas 3 y 10 son las más conflictivas. Con estas observaciones motivan a realizar un análisis de esas rutas, como el continente de destino que tengan en común, las navieras que operan por tales rutas, hasta el desempeño de la persona de Chartering encargada de la asignación de entregas en esas rutas.

| Ruta original ----- Nueva Ruta | Ruta 1 | Ruta 2 | Ruta 3 | Ruta 4 | Ruta 5 | Ruta 6 | Ruta 7 | Ruta 8 | Ruta 9 | Ruta 10 | Ruta 11 | Ruta 12 | Ruta 13 | Ruta 14 | Ruta 15 | Ruta 16 | Ruta 17 | Ruta 18 | Ruta 19 | Ruta 20 | Total general |
|--------------------------------------|----------|----------|------------|----------|-----------|-----------|-----------|-----------|-----------|------------|-----------|-----------|-----------|-----------|-----------|----------|-----------|-----------|-----------|-----------|---------------|
| Ruta 1 | | | | | 10 | | | | | | | | | | | | 10 | | | | 20 |
| Ruta 2 | | | | 2 | 30 | | | | | | | | | | | | 10 | 15 | | | 57 |
| Ruta 3 | 1 | 4 | 54 | 3 | 30 | | | | | | | 6 | | | | | 33 | 2 | 10 | | 143 |
| Ruta 4 | | | 39 | | | | | | | | | | | | | | 1 | | | | 40 |
| Ruta 5 | | | 3 | | | | | | | | 30 | | | | | | 1 | | | | 34 |
| Ruta 6 | | | | | | 1 | | 18 | | 1 | | | | | | | | | | | 20 |
| Ruta 7 | | | | | | | 11 | 3 | | 47 | | | | | | | | | | | 61 |
| Ruta 8 | | | | | | | | 18 | | 24 | | | | | | | | | | | 42 |
| Ruta 9 | | | | | | | | | | | | | | | | | | | | | 0 |
| Ruta 10 | | | | | | 38 | 1 | 53 | | 47 | | | | | | | | | | 16 | 155 |
| Ruta 11 | | | 10 | | | | | | | | | 4 | | | | | 10 | | | | 24 |
| Ruta 12 | | | | | | | | | | | | | 4 | | | | | | | | 0 |
| Ruta 13 | | | | | | | | | | | | | 1 | 22 | | | | | | 8 | 31 |
| Ruta 14 | | | | | | | | | | | | | 23 | | | | | | | | 23 |
| Ruta 15 | | | | | | | | | | | | | | | 59 | 4 | | | | | 63 |
| Ruta 16 | | | | | | | | | | | 2 | 6 | | | | | | | | | 8 |
| Ruta 17 | | | | | | | | 1 | | 25 | | | | | | | | | | | 26 |
| Ruta 18 | | | | | | | | | | | 10 | | | | | | | 15 | | | 25 |
| Ruta 19 | | | 10 | | | | | | 16 | | | | | | | | | | 18 | | 44 |
| Ruta 20 | | | | | | | | | | | | | 25 | 24 | | | | | | | 49 |
| Total general | 1 | 4 | 116 | 5 | 70 | 39 | 12 | 93 | 16 | 144 | 42 | 16 | 49 | 46 | 59 | 4 | 65 | 32 | 28 | 24 | 865 |

Tabla N°8: Flujo de reasignaciones entre rutas marítimas.

Fuente: Elaboración propia.

XII.V ANÁLISIS DE SENSIBILIDAD

La empresa debe negociar los contratos con las empresas navieras, estos contratos definen el precio o flete, y la cantidad de contenedores que Arauco deberá embarcar con tal naviera. Los contratos suelen durar uno, dos o cinco años, en esta línea, es interesante tener una forma de evaluar qué es lo más conveniente para la red Logística de Arauco.

Se propone utilizar el optimizador construido en esta memoria para hacer un análisis de sensibilidad de los parámetros de flete y capacidad de las naves sobre el costo total logístico y el nivel de servicio.

En la tabla N°9 se muestran los resultados de aumentar el flete y la capacidad de las naves un 10% para tres navieras por separado, es decir, se aumentaron los costos de Flete de una naviera un 10%, y manteniendo todo el resto con su configuración original, se ejecuta el optimizador para los 6 escenarios. Se procede de la misma manera con la capacidad de las naves.

Se observa que hay Navieras que son sensibles tanto en el costo logístico como en el nivel de servicio total, entre las Navieras analizadas, la naviera B es la más sensible al Flete, esto puede deberse a que esta Naviera va a puertos que ninguna otra Naviera puede ir, entonces un alza del 10% en el Flete impacta directamente en el costo logístico total. Por otro lado, puede que la Naviera B sea la más barata a algunos destinos, pero después del aumento del 10% en el Flete deja de serlo, entonces el optimizador prefiere otras naves que son más baratas, pero empeoran el nivel de servicio.

Hay que recordar que el proceso de optimización primero minimiza el costo logístico total y luego maximiza el nivel de servicio.

| Naviera | Variación Parámetros | Promedio Variación Costo Logístico [USD] | Promedio Variación Nivel de servicio |
|---------|----------------------|------------------------------------------|--------------------------------------|
| A | +10% Flete | \$ 151.580 | -0,5% |
| B | +10% Flete | \$ 312.200 | -1,3% |
| C | +10% Flete | \$ 34.660 | 0,0% |
| A | +10% Capacidad Naves | \$ -5.570 | 1,0% |
| B | +10% Capacidad Naves | \$ -29.300 | -0,8% |
| C | +10% Capacidad Naves | \$ - | 0,0% |

Tabla N°9: Variación promedio del costo logístico y nivel de servicio ante una variación en el Flete y Capacidad de las naves de cada Naviera.

Fuente: Elaboración propia.

La cantidad de combinaciones que se pueden analizar, respecto a los parámetros de Flete y capacidad, son muchas, en los resultados de la tabla N°9 solo muestran una variación de un 10% en estos parámetros de forma independiente y por Naviera. Pero en la negociación

de los contratos, el Flete y la Capacidad se pueden negociar por continente, ruta marítima, país o incluso puerto de destino, entonces la apertura por Naviera no es suficiente.

Supongamos que el Flete puede variar entre un 80% y un 120% con saltos de un 1% a la vez, y la Capacidad de las naves entre un 50% y un 200% con saltos de un 5% a la vez, en promedio una naviera va a 80 puertos distintos, en 30 países, 4 continentes y 3 rutas marítimas, finalmente, considerando que 10 escenarios (Instante en el tiempo) pueden ejemplificar la operación de un año completo, nos da un total de 340.56 millones de combinaciones solo para una Naviera. Si se quisiera evaluar más de dos Navieras al mismo tiempo, la cantidad de combinaciones serían de orden exponencial en el número de Navieras.

XII. CONCLUSIONES Y RECOMENDACIONES

En este trabajo se trató el problema asociado a la asignación de una entrega a un buque portacontenedores de una empresa forestal, definiendo para ello un modelo de datos robusto con el cual diseñar e implementar un optimizador de asignación, cuyo objetivo es alcanzar de la mejor forma posible el nivel de servicio meta al cliente, sujeto a un máximo costo.

- ✓ **Se diseño, construyo e implemento un Modelo de Datos para el Optimizador de asignación.** Cuyo objetivo es contar con datos robustos, limpios y consistentes que aseguren que el resultado entregado por el optimizador sea confiable. Se investigo sobre las herramientas existentes en el mercado y cuáles de ellas permiten una mayor probabilidad de éxito en la implementación en el equipo de Chartering.
- ✓ **Se estudio y reviso el estado del arte en técnicas de minería de datos para la estimación del parámetro asociado a los costos logísticos terrestres de una entrega.** La investigación fue documentada en el capítulo 9 de este documento, donde se describe tanto la metodología escogida, como los principales conceptos y algoritmos asociados al desarrollo de esta memoria.
- ✓ **Se determino un ahorro potencial de \$600.000 USD anual manteniendo el mismo nivel de servicio, mediante el desarrollo de un Modelo de Optimización, configurándolo para que minimice el costo logístico total para alcanzar un cierto nivel de servicio.** La estimación de este ahorro se construye a través de los resultados de distintos escenarios del año 2019.
- ✓ **Se determino la forma de calcular cuánto cuesta aumentar el nivel de servicio, el cual provee de información relevante en la toma de decisiones de la logístico de la empresa.** Es de responsabilidad del área de Chartering la decisión de cuánto dinero invertir en mejorar el nivel de servicio de Arauco, pero para tomar esta decisión, primero se debe conocer la base de cuál es el máximo nivel de servicio alcanzable al mínimo costo posible, información que hora puede ser conocida gracias a la herramienta desarrollada.

Uno de los aspectos a destacar del uso de cualquier herramienta de optimización es que debe tomarse como un apoyo para los tomadores de decisión, funcionando como una calculadora de KPI's complejos que complementan el conocimiento que ellos tienen sobre el negocio. Los resultados de los modelos de optimización, antes de ser aplicados, deben ser validados por los expertos del negocio.

Se ha logrado desarrollar e implementar con éxito una herramienta de asignación para el área de Chartering de la forestal Arauco, el cual consiste en un sistema de información, basado en un Modelo de Datos, un algoritmo de minería de datos, encargado de estimar

el parámetro de costos terrestres, y un modelo de optimización que se encarga de hacer la mejor asignación posible. A través de la herramienta es posible generar nuevo conocimiento para el negocio y apoyar la labor de los tomadores de decisión, permitiendo resolver el complejo problema de asignación de entregas a buques de la mejor forma posible.

Hasta ahora, el desarrollo de una herramienta de optimización robusta, cuyo resultado sea confiable, no se había logrado desarrollar de forma precisa dentro de la empresa. En este trabajo se entregan valores razonables del mínimo gasto logístico posible o la inversión necesaria para mejorar el nivel de servicio entregado al cliente, a través de un optimizador desarrollado por capas, las cuales permiten que las penalidades de costo y nivel de servicio sean independientes entre sí.

Un punto que considerar como mejora a futuro es el tema del respaldo de la información para generar indicadores estadísticamente representativos, utilizando la herramienta de optimización para evaluar diversos escenarios. Esto no fue considerado dentro de los alcances de este trabajo, sin embargo, es importante contar con alguna medida de recuperación de resultados históricos para su análisis.

Respecto al optimizador, se recomienda agregar una tercera etapa, en donde se intente minimizar la cantidad de reasignaciones de entregas. Este desarrollo se puede llevar a cabo implementando otra capa al modelo de optimización o iterar el modelo ya construido, restringiendo la cantidad de reasignaciones.

XIII. BIBLIOGRAFÍA

- [1] Arauco. Memoria Anual 2017 [en línea] <https://www.arauco.cl/peru/wpcontent/uploads/sites/22/2017/07/MEMORIA_ARAUCO_2017_FINAL.pdf> [Consulta: 2019/05/24]
- [2] William S. Kane, The truth about thriving in change, Financial Times Prentice Hall edición 1, 2008.
- [3] Knime analytic platform [en línea] <<https://www.knime.com/knime-software>> [Consulta: 2019/05/24]
- [3] Trevor Hastie, Robert Tibshirani y Jerome Friedman. The Elements of Statistical Learning. Second edition, 2001.
- [4] Shmueli, G., Patel, N. R., Bruce, P. C. Data mining for bussiness inteligence. Second edition. John Wiley and Sons, Hoboken, New Jersey. 2010.
- [5] Pottle, B., Agarwal, A. Predictive Analytics using Oracle Data mining. Data mining course. Student Guide, Edition 1.0. United State: Aishwarya Menom & Nikita Abraham 2016.
- [6] Henderson, J. C., Venkatraman, N. Estrategic alignment: Leveraging information Tecnology for Transforming Organizations. Lbm Systems Journal 1993.
- [7] Hong, S. J., Weiss, S. M., Advances in Predictive Model Generation for Data Mining (1999). [Recuperado de] <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.5584>>
- [8] Alina, F., Oscar, Y. B. S., Rodrigo, B. A. Introducción a la programación lineal, primera edición, CESA- Colegio de estudios superiores de administración, 2011.
- [9] Ashley, A., Optimziation Techniques in Coal Market: A global Costo Minimization And a Multi-Stage Procurement Strategy (2016).
- [10] Martin, S., Maritime Economics, Third Edition. Routledge 2008.
- [11] Breiman, L. 1996. "Bagging predictors". Machine Learning, Springer US.
- [12] U. M. FAYYAD, G. PIATETSKY-SHAPIRO and P. SMYTH. 1996. From data mining to knowledge discovery in Databases: an overview. Ai Magazine.

[13] L. ROKACH and O. Z. MAIMON. 2008. Data Mining with Decision Trees: Theory and Applications. World Scientific Publishing.

[14] TWO CROWS CORPORATION. 1999. Introduction to Data Mining and Knowledge Discovery. 3° ed. Two Crows Corporation.